

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTTTTATGCATTG
CTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAA
TTAAACAAGGCTGGCCTTGACTTGGAATACTGAGTTGTTTAGGACTTTCTATTGTGGCAAACCT
CCAGAAAACAACCCCTTTTGTGCTGCACATGTAAGTGGAGCTGTGCTTACCTTTGGTATGGGCTCAT
TATATATGTTTGTTCAGACCATCCTTTCCTACCAAATGCAGCCCAAATCCATGGCAAACAAGTC
TTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCAGTCTAGCATGCTGACTTGCTC
ATCAGTTTTGCACAGTGGCAATTTTGGGACTGATTTAGAACAGAACTCCATTGGAACCCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCAATTTTCCTTCTTT
GGTTTTTTCCTGACTTACATTCGTGATTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACA
TGGATTAACCCTCTATGACACTGCACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCA
GAGATATTTGATGAAAGGATAAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGG
TTCACAGAAGTTGCTTATTCTTCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCAT
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIIFS YITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFWIRLLLVWCGVSALSMLTCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFFLT YIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGAGTCCCGGCTGCAGCACCTGGGAGAAGGCAGACC
GTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAGGCAGGAGCCTTC
CTTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCTCCCAGATACTATTTTTTG
GATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTGAGTATGTTGTACAG
GTGATCTTCTCCGTGACGTTTGCATTTTCTTGACCATGTTTGAGCTCATCATCTTTGAAATCTTAGG
AGTATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGG
TTTTCATGGTGCCTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGA
CTGCTTTTTTCTGTCTCTTATGGCTGACCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCC
CATTCTCAGCCCAAAACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTTGGTGTGATTGGAG
TGACTCTCATGGCTCTTCTTTCTGGATTGGTGCTGTCAACTGCCCATACACTTACATGTCTTACTTC
CTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATATGAT
CATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAAGTGCATAACA
AACCATCAGGTTTCTGGGGAATGATAAAAAGTGTACCACCTCAGCATCAGGAAGTGAATCTTACT
CTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGCTTTTTCTGGAAACAGCTGATCT
ATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTCAAGGGGAAATATTTTAAATTTCTTGTT
ACTTTTTCTCTATTTACTGTGTTTGGAATTTTTCATGGCTACCATCAATATTGTTTTGATCGAGTT
GGGAAAACGGATCCTGTCAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGT
GAAGTTTGGTCCCAACATTTCTTCTTCTGTTGGAATAATCATCGTCACATCCATCAGAGGAT
TGCTGATCACTCTTACCAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCTCCAATGTCATTGTCCTG
CTATTAGCACAGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTT
AGAATACCGCACCATATACTCACTGAAGTCTTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTTG
ATGTGATCTTCTGGTCAGCGCTCTCTCTAGCATACTCTTCTCTATTTGGCTCACAAACAGGCACCA
GAGAAGCAATGGCACCTTGAACCTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTCAAATTTA
GATATAAGAGGGGGGAAAATGGAACAGGGCCTGACATTTTATAAAACAAACAAAATGCTATGGTAGC
ATTTTTACCTTCATAGCATACTCCTTCCCCGTCAGGTGATACTATGACCATGAGTAGCATCAGCCAG
AACATGAGAGGGAGAATACTCAAGACAATACTCAGCAGAGAGCATCCCGTGTGGATATGAGGCTGG
TGTAGAGGCGGAGAGGAGCAAGAACTAAAGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGT
CTATGGTAGCTGAGCCAAACACGTAGGATTTCCGTTTAAAGGTTTACATGGAAAAGGTTATAGCTTTG
CCTTGAGATTGACTCATTAAATCAGAGACTGTAACAAAAAAGGGCGGCCGCG
ACTCTAGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTATTGTCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQILFFGFGWLFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRYPFHWMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWKLGDP
FPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDILALERLLQ
TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQQEVDALAEELSRQ
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTDVPVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLLAQIMGY
FVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTTCC
 AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATCTGAGGT
 GTTTCCCTGGCTCTGAAGGGGTAGGCACGATGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT
 TCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCTTTCCATCCAGGT
 GTCATGCAGAATTATGGGGATCACCTTGTGAGCAAAAAGGCGAACCAGCAGCTGAATTTACAG
 AAGCTAAGGAGGCTGTAGGCTGCTGGGACTAAGTTTGGCCGGAAGGACCAAGTTGAAACAGCC
 TTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTTGGAGATGGATTCTGTTGCTCTCTAG
 GATTAGCCCAAACCCCAAGTGTGGGAAAAATGGGGTGGGTGTCCTGATTTGGAAGGTTCCAGTGA
 GCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGATACTTGGACTAACTCGTGCATTCCAGAA
 ATTATCACCACCAAAGATCCCATATTCAACACTCAAAGTCAACACAAACAACAAGATTTATGT
 CAGTGACAGTACCTACTCGGTGGCATCCCTTACTCTACAATACCTGCCCTACTACTCCTC
 CTGCTCCAGCTTCCACTTCTATTCCACGGAGAAAAAATTGATTTGTGTACAGAAGTTTTTATG
 GAACTAGCACCATGTCTACAGAACTGAACCATTTGTTGAAAATAAAGCAGCATTCAAGAATGA
 AGCTGCTGGGTTTGGAGGTGTCCCCACGGCTCTGCTAGTGCTTGTCTCCTCTTCTTTGGTGTG
 CAGCTGGTCTTGGATTTTCTATGTCAAAAGGTATGTGAAGGCCCTTCCCTTTTACAAACAAGAA
 CAGCAGAAGGAAATGATCGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCTAA
 TGAGGAATCAAAGAAACTGATAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAACTACCGTGC
 GATGCTGGAGCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTTCTTT
 CATGCTCCTTACCCTGCCCGTGGGGAAATCAAAGGGCCAAAGAACCAAGAAGAAAGTCCA
 CCCTTGGTTCCCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAA
 GCCCTTCTCCTTATTGTAACCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCCT
 TTCTAGCCTGGCTATGTCTTAATAATATCCCACTGGGAGAAAGGAGTTTGCAAAGTGAAGGAC
 CTAACATCTCATCAGTATCCAGTGGTAAAAAGGCCCTCCTGGCTGTCTGAGGCTAGGTGGGTTG
 AAAGCCAAGGAGTCACTGAGACCAAGGCTTCTCTACTGATTCCGCAGCTCAGACCTTTCTTCA
 GCTCTGAAAGAGAAACACGTATCCACCTGACATGTCTTCTGAGCCCGGTAAGAGCAAAAGAA
 GGCAGAAAAGTTTAGCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTCTGTAAA
 GCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACTGTCAGCAGGGACTGTAAAC
 ACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGAATCACTGTTTAGAACACACACA
 CTTACTTTTTCTGGTCTCTACCACTGCTGATATTTCTCTAGGAAATATACTTTTACAAGTAACA
 AAAATAAAACTCTTATAAATTCTATTTTTATCTGAGTTACAGAAATGATTACTAAGGAAGATT
 ACTCAGTAATTTGTTTAAAAAGTAATAAAATTCAACAAACATTTGCTGAATAGCTACTATATGTC
 AAGTGTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCCTCAAAAAATTGCACATAGTAG
 AACGCTATCTGGGAAGCTATTTTTTTCAGTTTTGATATTCTAGCTTATCTACTTCCAACTAAT
 TTTTATTTTTGCTGAGACTAATCTTATTCATTTTCTCTAATATGGCAACCATTATAACCTTAAT
 TATTATTAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAAGCACATTTTAAAGTGCC
 ATTAACAAATGTATCACTAGCCCTCCTTTTCCAACAAGAAGGACTGAGAGATGCAGAAATATT
 TGTGACAAAAAATTAAAGCATTTAGAAAACCTT

FIGURE 6

MARCFSLVLLLTISIWTRRLVQGSIRAEELSIQVSCRIMGITLVSKKANQQLNFTAKEACRLLG
LSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYN
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVK
RYVKAFPFNTNKNQKEMIETKVVKEEKANDSNPNNEESKKTDKNPNEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGCACCCGCGGGCCCCACCGCGCCGCTCCCGCATCTGCACCCGAGCCCGGC
GGCCTCCCGGGCGGAGCGAGCAGATCCAGTCCGGCCCCGAGCGCAACTCGGTCCAGTCCGGGGCGG
CGGCTGCGGGGCGCAGAGCGGAGATGCGAGCGGCTTGGGGCCACCCTGCTGTGCCTGCTGTGGCGG
CGGCGGTCCCCACGGCCCCCGCGCCGCTCCGACGGCGACCTCGGCTCCAGTCAAGCCCGGCCCG
GCTCTCAGCTACCCGAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGAGGTTGAGGAAGTATGAT
GGAGGACACGACGACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA
AAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAACAGAC
TGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGACGAAGAAGGCAGAAGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGGCCAGCATGTACTGCCAGTTTGCCAGCTTCCAGTAC
ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAGTGTGTGGAGACCA
GCTGTGTGTCTGGGGTCACTGCACCAAAATGGCCACCAGGGGCGAGCAATGGGACCATCTGTGACA
ACCAGAGGGACTGCCAGCCGGGGCTGTGCTGTGCCTTCAGAGAGGCTGCTGTTCCCTGTGTGC
ACACCCCTGCCCGTGGAGGGCGAGCTTTGCCATGACCCGCCAGCCGGCTTCTGGACCTCATCAC
CTGGGAGCTAGAGCCTGATGGAGCCTTGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGC
CCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGG
GAGATCCTGCTGCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTG
CCGCCGCTGCACTGCTGGGAGGGGAAGAGATTAGATCTGGACCAGGCTGTGGGTAGATGTGCAA
TAGAAATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCCTA
CATCTTCTTCCAGTAAGTTTCCCTCTGGCTTGACAGCATGAGGTGTGTGCAATTTGTTTACGCT
CCCCAGGCTGTTCTCCAGGCTTACAGTCTGGTGCTTGGGAGAGTCAGGCAGGGTTAAACTGCA
GGAGCAGTTTGGCCACCCCTGTCCAGATTATTTGGCTGCTTTGCCTCTACCAAGTTGGCAGACAGCCG
TTTGTCTACATGGCTTTTGATAATTGTTTGGGGGAGGAGATGGAACAATGTGGAGTCTCCCTC
TGATTGGTTTTTGGGGAATGTGGAGAAGAGTGCCCTGCTTTGCAAACATCAACCTGGCAAAATG
CAACAAATGAATTTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTGTGCCTTCAGCTGTTGC
AGATGAAATGTTCTGTTACCCCTGCATTACATGTGTTTATTCATCCAGCAGTGTGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTTCATATCCAAGATCAATTCCCTCTCTCAGCACAGCCTGGGG
AGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCATCTGGTTGTGACTCTAAGCTCAGTGCTCT
CTCCACTACCCACACAGCCTTGGTGCCACCAAAAGTGCTCCCCAAAAGGAAGGAGAATGGGAT
TTTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAACAAATCTCACATCCCTCTAAAAGTAAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGAT
ATTGACACTGTCCCTCTTTGGCAGTTGCATTAGTAACCTTGAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACCTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGC
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAATGATG
TTTTCAGGTGTATGGACTGTTGCCACCATGTATTATCCAGAGTCTTAAAGTTTAAAGTTGCA
CATGATTGTATAAGCATGCTTTCTTTGAGTTTTAAATTATGTATAAACATAAGTTGCATTTAGAA
ATCAAGCATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLLAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTGQMVFE
TVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRMMLCTRDECCGDQLCVWGHC
TKMATRGSNGTICDNQRDCQPGGCCAFQRGLLFPVCTPLPVEGELCHDPASRLDLITWELEPDG
ALDRPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQELEDLE
RSLTEEMALGEPAAAAALLGEEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCAC
GGCCCACCTTGTGAACCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCATCCAAAG
GCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCTGGGGCTCTTCTGGACCCTT
AACTGGGTACTGGCCCTGGGCCAATGCGTCTCGCTGGAGCCTTTCCTCCTTCTACTGGGCCTT
CCACAAGCCCCAGGACATCCCTACCTTCCCTTAATCTCTGCCTTCATCCGCACACTCCGTACC
ACACTGGGTCAATTGGCATTGGAGCCCTCATCTGACCCTTGTGCAGATAGCCGGGTCACTTG
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGCCGCTGCATCATGTGCTGTTT
CAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCCTAAACCGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAAAATGCGTTCATGCTACTCATGCGAAC
ATTGTCAGGGTGGTCTGCTGGACAAAGTCACAGACCTGCTGCTGTTCTTTGGGAAGCTGCTGGT
GGTCGGAGGCGTGGGGTCTGTCTTCTTTTTTCTCCGGTCGCATCCCGGGCTGGGTAAAG
ACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCCATCATGACCTCCATCCTGGGGGCCTAT
GTCATCGCCAGCGCTTCTTCAGCGTTTTTCGGCATGTGTGTGGACACGCTCTTCTCTGCTTCCT
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGCCCTACTACATGTCCAAGAGCCTTCTAA
AGATTCTGGGCAAGAAGAACCAGGGCGCCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG
CCCTGATCCAGGACTGCACCCACCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGT
CTCCATTTTGTGGTAAAAAAGGTTTTAGGCCAGGCGCGTGGCTCACGCCTGTAATCCAACACT
TTGAGAGGCTGAGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTG
AAACCTCCGCTCTTATTAATAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCA
GCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGA
GATCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAAACAA
AAGATTTTATTAAAGATATTTTGTTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLEWTL
NWVLALGQCVLAFASFYWAFHKPDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN
IVRVVVLDKVTDLLFFGKLLVVGGVLSFFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY
VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GCCCCGCGCCCGGCGCGGGCGCCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGAGCCTGC
TCCCTGCTCAGCTGCGCTCCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAGCTGCTGCCCCGC
CAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCACGTTCTTCTCTTCTGCGGGTGTGGTGTCCA
TCATTATGCTGAGCCCGGGCGTGGAGAGTCAGCTCTACAAGCTGCCCTGGGTGTGTGAGGAGGGGGCC
GGGATCCCCACCGTCTGCAGGGCCACATCGACTGTGGCTCCCTGCTTGGCTACCGCGCTGTCTACCG
CATGTGCTTCGCCACGGCGGCTTCTTCTTCTTCTTTTACCCTGCTCATGCTCTGCGTGAGCAGCA
GCCGGGACCCCGGGCTGCCATCCAGAATGGGTTTGGTTCTTTAAGTTCCTGATCCTGGTGGGCCTC
ACCGTGGGTGCCTTCTACATCCCTGACGGCTCCTTCACCAACATCTGGTTCTACTTCGGCGTGTGGG
CTCCTTCTTCTTCTCATCCAGCTGGTGTCTCATCGACTTTGCGCACTCCTGGAACCAGCGGT
GGCTGGGCAAGGCCGAGGAGTGCGATTCCCGTGCCTGGTACGCGAGGCCTCTTCTTCTTCTCTCTC
TTCTACTTGCTGTGATCGCGGCCGTGGCGCTGATGTTTATGTACTTACTGAGCCAGCGGCTGCCA
CGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCTGTCTCTGC
CCAAGTCCAGGACGCCCAGCCCAACTCGGTCTGCTGCAGGCCTCGGTTCATCACCTCTACACCATG
TTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCATTTGCCAACCCAGCT
GGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAGACCCAGTGGTGGGATGCCCGAGCATTG
TGGGCCTCATCATCTTCTCTGTGACCCCTCTTCATCAGTCTGCGCTCCTCAGACCACCGGCAGGTG
AACAGCCTGATGCAGACCGAGGAGTGCCCACTATGCTAGACGCCACACAGCAGCAGCAGCAGGT
GGCAGCCTGTGAGGGCCGGGCTTTGACAACGAGCAGGACGGCGTCACCTACAGTACTCCTTCTTCC
ACTTCTGCTGGTGTGGCCTCACTGCACGTCATGATGACGCTCACCAACTGGTACAAGCCCGGTGAG
ACCCGGAAGATGATCAGCACGTGGACCGCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCT
CCTCTACCTGTGGACCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTGAGGCGAGCCTCA
CAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCTGCCCCCTC
CCCCACCAATCAGCCAGGCTGAGCCCCCACCCTGCCCCAGCTCCAGGACCTGCCCCTGAGCCGGGC
CTTCTAGTCGTAGTGCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCCCGCCAC
ACCCACACGGTGGAGCTGCCTCTTCTTCCCTCCTCCTGTTGCCATACTCAGCATCTCGGATGAA
AGGGCTCCCTTGTCTCAGGCTCCACGGGAGCGGGGCTGCTGGAGAGAGCGGGGAATCCACCACAG
TGGGGCATCCGGCACTGAAGCCCTGGTGTCTCTGGTACGTCCCCAGGGGACCCTGCCCCCTTCTCTG
GACTTCGTGCCTTACTGAGTCTCTAAGACTTTTCTAATAAACAAGCCAGTGCGGTGTAACAAAAA

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFFLGLVLSIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSLFILIQLVLLIDFAHSWNQRWLKAE
ECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHGKVFISLNLTFVCVVSIAAVLPKV
QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQVAACEGRAFDNEQDGVITYSY
SFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVVWKICASWAGLLLYLWTLVAPLLLRNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGGCCGGCCAGGAACCACCCGTTAAGGTGTCTTCTCTTTAGGGATGGTGA
GGTTGGAAAAAGACTCCTGTAACCCCTCCTCCAGGATGAACCACCTGCCAGAAGACATGGAGAACG
CTCTCACCGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCACACAA
CTCATGGCCAGGATTGAGTCTTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTTCTGTTTGTGTTGTACCTTTGACCTCTTATTCTGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTTCATAT
TTTGATATATTTCTTCTGGCAGTTTTTCGATTTAAAGTGTTAATACTTGCATATGCTGTGTGCAG
ACTGCGCCATTGGTGGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTTTACTAGCAAAAG
TGATCCTTTTCGAAGCTTTTCTCTCAAGGGGCTTTTGCTATGTGCTGCCCATCATTTTCATTCATC
CTTGCTGGATTGAGACGTGGTTTCCTGGATTTCAAAGTGTTACCTCAAGAAGCAGAAGAAGAAAA
CAGACTCCTGATAGTTCAGGATGCTTCAGAGAGGGCAGCACTTATACCTGGTGGTCTTTCTGATG
GTCAGTTTTTATCCCTCCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT
GAGAAACCACTTTTAGAACTATGAGTACTACTTTTGTTAAATGTGAAAAACCTCACAGAAAGTC
ATCGAGGCAAAAAGAGGCAGGCAGTGGAGTCTCCCTGTCGACAGTAAAGTTGAAATGGTGACGTC
CACTGCTGGCTTTATTGAACAGCTAATAAAGATTTATTTATTGTAATACCTCACAAACGTTGTAC
CATATCCATGCACATTTAGTTGCCTGCCTGTGGCTGGTAAGGTAATGTCATGATTCATCCTCTCT
TCAGTGAGACTGAGCCTGATGTGTTAACAATAAGGTGAAGAAAGTCTGTGCTGTATTCTTAATC
AAAAGACTTAATATATTGAAGTAACACTTTTTTAGTAAGCAAGATACCTTTTTATTTCATTCAC
AGAATGGAATTTTTTGTTCATGTCTCAGATTTATTTTGTATTTCTTTTTTAACACTCTACATT
TCCCTTGTTTTTTAACCTCATGCACATGTGCTCTTTGTACAGTTTTTAAAAAGTGAATAAAATCTG
ACATGTCAATGTGGCTAGTTTTATTTTCTGTTTTGCATTATGTGTATGGCCTGAAGTGTGGA
CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAAAATGTCACCAGACATTTGTATTATTT
TTATCATGAAATCATGTTTTTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTTTGAATGC
ACAAAATGACTTAAACCATTATATCATGTTTCCTTTGCGTTTCAGCCAATTTCAATTAAAATGAA
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVLQPQAEENRLLIVQDASER
AALIPGGGLSDGQFYSPPESEAGSEEAEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCCGACCCGCCAGGAAAGACTGAGG
CCGCGGCCTGCCCCGCGGCTCCCTGCGCCGCGCCGCTCCCGGGACAGAAGATGTGCTCCAG
GGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCTGGGGTGCAGGGTGGCCAT
CCGGCTGCCAGTGCCAGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGGACCAGGTGCCC
CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCATCACCATGCTCGACGC
AGGCAGCTTTGCCGGCCTGCCGGGCTGCAGCTCCTGGACCTGTCACAGAACCAGATCGCCAGCC
TGCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTGGACCTGACGGCCAACAGGCTG
CATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCTCGAGCGCCTTACCTGGGCAAGAA
CCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACCGCCTCCTGGAGCTCAAGCTGC
AGGACAACGAGCTGCGGGCACTGCCCGGCTGCGCCTGCCCGCCTGCTGCTGCTGGACCTCAGC
CACAAACAGCTCCTGGCCCTGGAGCCCGCATCCTGGACACTGCCAACGTGGAGGCGCTGCGGCT
GGCTGGTCTGGGGCTGCAGCAGCTGGACGAGGGGCTCTTACGCGCTTGCACACCTCCACGACC
TGGATGTGTCCGACAACCAGCTGGAGCGAGTGCCACCTGTGATCCGAGGCTCCGGGGCTGACG
CGCCTGCGGCTGGCCGGCAACCCCGATTGCCAGCTGCGGGCCGAGGACCTGGCCGGCTGGC
TGCCCTGCAAGAGCTGGATGTGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCC
TCTTCCCCCGCTGCGGCTGCTGGCAGCTGCCCGCAACCCCTTCAACTGCGTGTGCCCCCTGAGC
TGTTTGGCCCTGGGTGCGCGAGAGCCAGTCACTGCGCAGCCCTGAGGAGACGCGCTGCCA
CTTCCCGCCCAAGAACGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCAG
CCACCACCACACAGCCACAGTGCCCAACGAGGCCCGTGGTGGGGAGGCCACAGCCTTGTCT
TCTAGCTTGGCTCCTACCTGGCTTAGCCCCACAGCGCCGCGCCACTGAGGCCCCAGCCCGCCCTC
CACTGCCCCACGACTGTAGGCGCTGTCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCTCA
ATGGGGGCACATGCCACCTGGGGACACGGCACACCTGGCGTGCTTGTGCCCCGAAGGCTTCACG
GGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCAGCCCTACACCAGTCAAGCCGAG
GCCACCACGGTCCCTGACCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGCGTGGGGCTGC
AGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTATCGCAACCTATCG
GGCCCTGATAAGCGGCTGGTGACGCTGCGACTGCCTGCCTCGCTCGCTGAGTACAGGTACCCCA
GCTGCGGCCCCAACGCCACTTACTCCGTCTGTGTATGCCTTTGGGGCCCGGGCGGGTGCCGGAGG
GCGAGGAGGCCCTGCGGGGAGGCCCATACCCCCAGCCGTCCACTCCAACCACGCCCCAGTCACC
CAGGCCCGCGAGGGCAACCTGCCGCTCCTCATTGCGCCCGCCCTGGCCGCGGTGCTCCTGGCCGC
GCTGGCTGCGGTGGGGGACGCTACTGTGTGCGGCGGGGCGGGCCATGGCAGCAGCGGCTCAGG
ACAAAGGGCAGGTGGGGCCAGGGCTGGGCCCTGGAAGTGGAGGGAGTGAAGTCCCCTTGAGG
CCAGGCCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGGTCTGAGTGTGAGGTGCC
ACTCATGGGCTTCCAGGGCTGGCCTCCAGTCAACCCCTCCACGCAAGCCCTACATCTAAGCCA
GAGAGAGACAGGGCAGCTGGGGCCGGGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCC
ACACCACGTAAGTTCTCAGTCCCAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCT
GGGCCCTGTTCCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCAGCTGACGAGCC
CTAACGTCCCCAGAACCAGTGCCTATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTC
CCTGGGCACGGCGGGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCTGCTGGGCTCTCCAC
TCCAGGCGGACCCTGGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAGGGAGAGCGGGTAGGC
GGCTGTGTGACTTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAACTGGAAAGGAAGATGC
TTTAGGAACATGTTTTGCTTTTTTAAATATATATATTTATAAGAGATCCTTTCCCATTTATTCT
GGGAAGATGTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTTGTAAGACAAACGATGATATGAA
GGCCTTTTGTAAAGAAAAATAAAAGATGAAGTGTGAAA

FIGURE 16

MCSRVP LLLPL LLL L L L L L L A L G P G V Q G C P S G C Q C S Q P Q T V F C T A R Q G T T V P R D V P P D T V G L Y V F E N G I T
M L D A G S F A G L P G L Q L L D L S Q N Q I A S L P S G V F Q P L A N L S N L D L T A N R L H E I T N E T F R G L R R L E R L Y
L G K N R I R H I Q P G A F D T L D R L L E L K L Q D N E L R A L P P L R L P R L L L L D L S H N S L L A L E P G I L D T A N V E
A L R L A G L G L Q Q L D E G L F S R L R N L H D L D V S D N Q L E R V P P V I R G L R G L T R L R L A G N T R I A Q L R P E D L
A G L A A L Q E L D V S N L S L Q A L P G D L S G L F P R L R L L A A A R N P F N C V C P L S W F G P W V R E S H V T L A S P E E
T R C H F P P K N A G R L L L E L D Y A D F G C P A T T T T A T V P T T R P V V R E P T A L S S S L A P T W L S P T A P A T E A P
S P P S T A P P T V G P V P Q P Q D C P P S T C L N G G T C H L G T R H H L A C L C P E G F T G L Y C E S Q M G Q G T R P S P T P
V T P R P P R S L T L G I E P V S P T S L R V G L Q R Y L Q G S S V Q L R S L R L T Y R N L S G P D K R L V T L R L P A S L A E Y
T V T Q L R P N A T Y S V C V M P L G P G R V P E G E E A C G E A H T P P A V H S N H A P V T Q A R E G N L P L L I A P A L A A V
L L A A L A A V G A A Y C V R R G R A M A A A A Q D K G Q V G P G A G P L E L E G V K V P L E P G P K A T E G G E A L P S G S E
C E V P L M G F P G P G L Q S P L H A K P Y I

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGGCGAGGCGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTCATGCG
GGTCCGGATAGGGCTGACGCTGCTGCTGTGTGCGGTGCTGCTGAGCTTGGCCTCGGCGTCTCGG
ATGAAGAAGGCAGCCAGGATGAATCCTTAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTTCTTGATTCAGAAGAATCTGA
ATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAAAAGTGTACAG
AAGATATCAGCTTTCTAGAGTCTCCAAATCCAGAAAACAAGGACTATGAAGAGCCAAAGAAAGTA
CGGAAACCAGCTTTGACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCT
TTTCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTG
CTACAACCTATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAAGTGAATGAAAATCCTTAATGGAAG
CAATAAGAAAAGCCAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTATATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATCCAG
GCAGCGAGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCCAAGGGACAGACTGCTCTTGG
CTTCTGTATGCCTCTGGACTTGGTGTTAATTCAAGTCAGGCAAAGGCTCTTGTATATTATACAT
TTGGAGCTCTTGGGGCAATCTAATAGCCCATGTTTGGTAAGTAGACTTTAGTGGAAGGCT
AATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACTTTTTCAGCTTTCATGATC
CAGATTTGCTTGTATTAGACCAAATATTCAGTTGAACTTCCTTCAAATCTTGTTAATGGATAT
AACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACAATTTTTCTTTAAATGATTAG
TTTGGCTGATTGCCCCATAAAAGAGAGATCTGATAAATGGCTCTTTTAAATTTTCTCTGAGTTG
GAATTGTCAGAATCATTTTTTACATTAGATTATCATAATTTTAAAAATTTTCTTTAGTTTTTCA
AAATTTTGTAATGGTGGCTATAGAAAACAACATGAAATATTATACAATTTTTGCAACAATGC
CCTAAGAATTGTTAAATTCATGGAGTTATTTGTGCAGAATGACTCCAGAGAGCTCTACTTTCTG
TTTTTTACTTTTCATGATTGGCTGTCTTCCCATTTATTCTGGTCATTTATTGCTAGTGACACTGT
GCCTGCTTCCAGTAGTCTCATTTTCCCTATTTTGCTAATTTGTTACTTTTTCTTTGCTAATTTGG
AAGATTAACATTTTTAATAAAATTATGTCTAAGATTAATAAAAAAAAAAAAAAAAAAAAAAAAAA
AA

FIGURE 18

MRVRIGLTLLLCVLLSLASASSDEEGSQDESLSKTTLTSDSVKDHTTAGRVVAGQIFLDSESEL
ESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHGEPCHFPFLFDK
EYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEBAKRRQMGEAEMMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERVSYALLFGDYLPQNIQAAREMFELTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYTFGALGGNLIHMLVLSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCAGATTTTAAGCCCATCTGCAGTGGAATTCATGAAGTAGCAAGAGGACACCATCTTCTT
GTATTATACAAGAAAGGAGTGACCTATCACACACAGGGGGAAAAATGCTCTTTTGGGTGCTAGG
CCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAAATAAGATTGAAGACATCACTG
ATAAGTACATTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCCAGAACTTTTGAT
AAAAAGGGATTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAACAGCTTTAAAGGCAGA
AACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG
CCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGTCTCTGGGGTCTGATCAATAATGCTGGTGTT
CCCGGCGTGCTGGCTCCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
CCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTTGGTCAAGAAAGCTCAAGGGAGAG
TTATTAATGTCTCCAGTGTGGAGGTCGCCTTGCAATCGTTGGAGGGGGCTATACTCCATCCAAA
TATGCAGTGAAGGTTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTTGGTGTGCACGTCTC
ATGCATTGAACCAGGATTGTTCAAAACAACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC
TCGCCATTTGGGAGCAGCTGTCTCCAGACATCAAACAACAATATGGAGAAGGTTACATTGAAAAA
AGTCTAGACAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTG
CATGGACCACGCTCTAACAAGTCTCTTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAAA
TTTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAA
GCAGAGCTGGCTAATCCCAAGGCAGTGTGACTCAGCTAACCACAAATGTCTCCTCCAGGCTATGA
AATTGGCCGATTTCAAGAACACATCTCCTTTCAACCCCATTCCTTATCTGCTCCACCTGGACT
CATTTAGATCGTGCTTATTTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGT
CCCTGCTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCTGCCCT
GTATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTTGCCCATTCAAAATGA
TCTTTACCGTGGCCTGCCCCATGCTTATGGTCCCCAGCATTTACAGTAACTTGTGAATGTTAAGT
ATCATCTCTTATCTAAATATTAAAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGLKIEDITDKYIFITGCDSGFGNLAARTFDKKG FHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNOVGEGKLWGLINNAGVPGVLAPTDWLTLEDY
REPIEVNLFGLISVTLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
AFGVHVSCEIEPGLFKTNLADPVKVEKKLAIWEQLSPDIKQQYGEGYIEKSLDKLKGKNSYVNMD
LSPVVECMDHALTSLEFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGCGGCGGTAGCATGGAGGGGAGAGTACGTGGGCGGTGCTCTCGGGCTTTGTGCTCGGCG
CACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAAGTAAAA
GGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTATACAATTGA
CATTAGAAATATATTCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAGGCGAAGTAAATG
AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATTCGGT
CGTCATTAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAACCTGCAGGAGCATT
TTCAAACCAAGACCTTGTTTTCTGCTATTAACACCAAGTATAATAACAGAAAGCTGCTCTACTC
ATCGACTGGAACATTCCCTATATAAACCTCAAAAAGGACTTTTTACAGGGTACCTTTAGTGGTT
GCCAATCTGGGCATGTCTGAACAACTGGGTATATAAACTGTATCAGGTTCCGTATGTCCACTGG
TTTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTTGAAGAAGATGGATCCTTAAAGGAGG
TACATAAGATAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAAGTG
GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAAACAGATTAAAACGAGAAATTGA
GAAAAGGAGAGGAGCACAGATTAGGCAGCAAGAGAGAAGAACATCCAAAAGACCTCAGGAGA
ACATTTTTCTTTGTCAGGCATTACGGACCTTTTTTCCAAATTCTGAATTTCTTCATTATGTGTT
ATGTCTTTAAAAAATAGACATGTTTCTAAAAGTAGCTGTAACACCAACCACCTCGATGTAGT
AGACAATCTGACCTTAATGGTAGAACACACTGACATTCTGAAGCTAGTCCAGCTAGTACACCAC
AAATCATTAAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTA
GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAT
GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTC
CTACATTTTGATCCTTTTAACCTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAAC
ATTTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTTACC
TGTTTGAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAACATCA
GATGCTTTTATTTCCAAACCTTTTTTTCACCTTTCACTAAGTTGTTGAGGGGAAGGCTTACACAG
ACACATTCTTTAGAAATTGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCAGCACT
TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATT
GAGACCATGTCTATTAATAAATAAATGAAAAGCAAGAATAGCCTTATTTCAAATATGGAAA
GAAATTTATATGAAAATTTATCTGAGTCATTAAATTTCTCCTTAAGTGATACTTTTTTAGAAGTA
CATTATGGCTAGAGTTGCCAGATAAATGCTGGATATCATGCAATAAATTTGCAAAACATCATCT
AAAATTTAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTSDTEGFLLEVKGEAKNSITDSQMDDVEVVYTIDIQKYI
PCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL
VFLLLTPSIITESCSTHRLEHSLYKPKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSSRAV
QTHSSKFFEEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVNRLKREIEKRRGA
QIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCYNHHLVDVNDLTL
MVEHTDIPEASPASTPQIIKHKALDLDLRWQFKRSRLDTQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTE

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGACGAGCGGACCAGCGCAGGGCAGCCCAA
GCAGCGCGCAGCGAACGCCCGCCGCCGCCACACCCTCTGCGGTCCCCGCGCGCCTGCCACCCCTCCCTCCTTCCCC
GCGTCCCCGCTCGCCGCGCAGTCAGCTTGCCGGGTTGCTGCCCGCGAAACCCGAGGTACACAGCCCGCGCCTCT
GCTTCCCTGGGCGCGCGCGCCTCCACGCCCTCCTTCTCCCTGGCCCGCGCCTGGCACCGGGACCGTTGCCTGA
CGCGAGGCCAGCTCTACTTTTCCGCCCGCCTCTCCTCGCCTGCTCGCCTCTTCCACCAACTCCAACCTCCTTCTCCC
TCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCGCTGCCGTAGCGCCGCTTCCCGTCCGGTCCCAAA
GGTGGGAACGCGTCCGCCCGGCCGACCATGGCACGGTTGCGCTTGCCCGCCTTCTCTGCACCTGGCAGTGCTC
AGCGCCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAGTTGCTCGGAAGTGCAGCTCTTTACGTGTCCAAGGCTTC
AACAGAACGATGCCCCCTCCACGAGATCAACGGTGATCATTTGAAGATCTGTCCCGAGGTTCTACCTGCTGCTCT
CAAGAGATGGAGGAGAAGTACAGCCTGCAAGTAAAGATGATTTCAAAAGTGTGGTCAGCGAACAGTGAATCATTTG
CAAGCTGTCTTTGCTTACGTTACAAGAAGTTTGATGAATCTTCAAAGAATACTTGAATAAGCAGAGAAATCCCTG
AATGATATGTTTGTGAAGACATATGGCCATTATACATGCAAAATTCTGAGCTATTTAAAGATCTCTTCGTAGAGTTG
AAACGTTACTACGTGGTGGAAATGTGAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCTGGAGCGGATG
TTCGCCCTGGTGAACCTCCAGTACCACTTTACAGATGATATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAG
CCCTTCGGAGATGTCCCTCGCAAATTGAAGCTCCAGGTTACTCGTGCTTTTGTAGCAGCCGTAATTTCCGCTCAAGC
TTAGCGGTGGCGGAGATGTCTGAGCAAGGTCTCCGTGGTAAACCCACAGCCAGTGTACCCATGCCCTGTTGAAG
ATGATCTACTGCTCCCACTGCCGGGCTCTCGTGACTGTGAAGCCATGTTACAATACTGCTCAAACATCATGAGAGGC
TGTTTGGCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTATAGATGCTATGCTGATGGTGGCAGAGAGGCTA
GAGGGTCCCTTTCAACATTGAATCGGTGATGATCCCATGATGTGAAGATTTCTGATGCTATTATGAACATGCAGGAT
AATAGTGTTCAGTGTCTCAGAAGGTTTCCAGGGATGTGGACCCCCCAAGCCCTCCAGCTGGACGAATTTCTCGT
TCCATCTCTGAAAGTGCTTCACTGCTCGCTTCAAGCACATCACCCGAGGAACGCCCAACACAGCAGCTGGCACT
AGTTTGGACCGACTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATTTCTGGTCTCCCTTCCGAGCAAC
GTTTGCACGATGAGAGGATGGCTGCAGGAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTAC
CTGTTGCACTGACAGGAAATGGATTAGCCAACAGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAAACAGAC
ATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGAAGAATGCATACAATGGGAACGACGTG
GACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGAAGTGGCTGTGAGTATCAGCAGTGCCCTTCAGAG
TTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGTGCTGCTCGCTCGGGCA
CAGGCCTACCTCCTCACTGTCTTGCATCTTGTCTGTTATGCAAGAGAGTGGAGATAATTCTCAAACCTCTGAG
AAAAAGTGTTCATCAAAAGTTAAAGGCACCACTTATCACTTTTCTACCATCCTAGTGACTTTGCTTTTAAATGAA
TGGACAACAATGTACAGTTTTTACTATGTGGCCACTGTTTAAAGAGTGTGACTTTGTTTTCTCATTCACTTTTGGG
AGGAAAAGGACTGTGCATTGAGTTGGTTCCTGCTCCCCAAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAA
CTATAGTTAGTTGTGCATTGTGATTTTATCACTCTATTATTTGTTGTATGTTTTTCTCATTTCGTTTGTGGGT
TTTTTTTCCAACGTGATCTCGCCTTGTTCCTTACAAGCAAACAGGGTCCCTTCTTGGCACGTAACATGTACGTATT
TCTGAATATTAATAGCTGTACAGAAGCAGGTTTTATTTATCATGTTATCTTATTAAGAAAAAGCCCAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQGST
CCSQEMEKEYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMFVKTYGH
LYMQNSELFKDLFVELKRYVVGNVNLEEMLNDFWARLLERMFRLVNSQYHFTDEYLECVSKYTE
QLKPFQDVPRKLLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHALLKMIYCSHCRGL
VTVKPCYNYCSNIMRGCLANQGDLDFEWNQFIDAMLMAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFQCGPPKPLPAGRISRSISESAFSARFRPHHPEERPTTAAGTSLDRLVTDVK
EKLKQAKKFWSSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGLANQGNNPEVQVDT
KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACT
TACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCACAGGGACCGAGGC
CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATTGGTTCTTGAG
AGCCCCGAGAAGAAAATTCTATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAAACAAGCATTCCAGA
GCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGTAGGAGCTCTG
AGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTC
TTCTTCTCCCACTCACTCTCCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCA
TGTTTTTCAAGATCATTTTGTGTTGCTCTCTCTAGTGTCTTCTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAACTGTAGCTTCCT
AGCTAGTGTCAATTAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTT
AAATGTCAAAAAAAAAAAAAAAAAA

FIGURE 26

MKVLISL L L L L L L P L M L M S M V S S S L N P G V A R G H R D R G Q A S R R W L Q E G G Q E C E C K D W F L R A P R R K F M
T V S G L P K K Q C P C D H F K G N V K K T R H Q R H H R K P N K H S R A C Q Q F L K Q C Q L R S F A L P L

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGACATCCTGGTCCCACCTCCTGCAGCTGCTGGTGCTGCTTCTTACCCTGCC
CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCTGTGCAAAAGCTACTTCCCCTACCTGA
TGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTCTTCAGCCAG
ATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGGCTGCGGAACCGGAGC
CAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCCACTTTGAGA
AGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTTGTGGTGGCTCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGTGTG
CTCTGTGCAGAGCCCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGC
TCTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCCTTCATGTGGCAGCAAGTTTTC
GAGCCCACCTGGAAACACATTGGGGATGGCTGCTGCCTCACCAGAGAGACCTGGAAGGATCTTGA
GAACGCCCAGTTCTCCGAAATCCAAATGGAACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTG
GGCCCCACATCATGGGAAGGCTGTCAAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCC
TTCCCCAGCCTCCAATTAGAACAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTTA
GCAGAATGAGAGAAGACATTTCATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGC
AATCTCTAACTTCAATCCCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCCTGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTC
CCAATGTTGTCCCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACAC
CCATGCGTCTCTAGGAACTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGAC
CCTCTCTCCCCACTACCACCTTCTTCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGAA
ACCACG

FIGURE 28

MDILVPLLQLLVLLLTPLPLHLMALLGCWQPLCKSYFPYLMVLTPKSNRKMESKKRELFSQIKGL
TGASGKVALLELGCGTGANFQFYPPGCRVTCCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDEVVCTLVLCVQSPRKVLQEVRRVLRPGGVLEFFWEHVAEPIYGSWAFMWQQVFPTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSFSSKALICSFPSL
QLEQATHQPIYLPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTTGCCTATCCACCTCCCCAAGCCCCTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCCAGAAAGTCTCTTG
CCACTGACGCCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTCCTGCCTCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGTGGGGGATGGCTA
AGAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGGCTGCCTTATTAAA
GTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGCCCTGTTCAATTAAGAAATT
GTTCCCTTCCCCTGTGTTCAATGTTTGTAAGATTGTTCTGTGTAAATATGTCTTTATAATAAAC
AGTTAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLLLLLLLLLKGSCLWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCSQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAAGTT
CCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
ATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGAATCATGTCGG
GAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCCATGATGTTTACC
TTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTCAATTGGTTATTTTGGGATTGTTGTT
TGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA
CAGAAAGGGAAAATATGAAGTGCCTGCTGGGGTTTGTCTATCGTATCCACAGGCATCACGGCAGTG
CTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTTCCAAATCAC
AAATAAAGCCATCAGCAGTGCTCCCTTCTGCTGTTCCAGCCACTGTGGACATTTGCCATCCTCA
TTTTCTTCTGGGTCCCTCTGGGTGGCTGTGCTGCTGAGCCTGGGAAGTGCAGGAGCTGCCAGGTT
ATGGAAGGCGGCCAAGTGAATATAAGCCCCCTTCGGGCATTCCGTACATGTGGTTCGTACCATTT
AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGGCTGCCAGCAAATGACTATAGCTGGGGCAG
TGGTTACTTGTATTTCAACAGAAGTAAAAATGATCCTCCTGATCATCCCATCCTTTTCGTCTCTC
TCCATTCTCTTCTTCTACCATCAAGGAACCGTTGTGAAAGGGTCATTTTTAATCTCTGTGGTGAG
GATTCCGAGAATCATTGTATGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGT
CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTC
AACCAGAATGCATATACTACAAGTCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGC
ATTCAAAATCTTGTCCAAGAACTCAAGTCACCTTACATCTATTAAGTCTTGGAGACTTCATAA
TTTTTCTAGGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACTAC
AATCGGGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTGCCTACTTAGTAGC
CCATAGTTTTTTTATCTGTGTTTGAAGTGTGCTGGATGCACTTTTCTGTGTTTGTCTGTGATC
TGGAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTA
AAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAAGGAATGAGGA
GGGAACAGAACTCCAGGCCATTGTGAGATAGATACCCATTTAGGTATCTGTACCTGGAAAACATT
TCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTTT
TTAAAAGACCTAATAAACCCCTATTCTTCCTCAAAA

FIGURE 32

MSGRDTILGLCILALALSLAMFTFRFITLLVHIFISLVILGLLFVCGVLWWLYDYTDNLSIE
LDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQPLWTF
ILIFFVVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQQTIA
GAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVMYMQLKEQQHG
ALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD
FIIFLGKVLVVCFTVFGGLMAFNYNRAQVWAVPLLVAFFAYLVAHSFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTTGGACCTCTCCCTGTTTCTTCTTAGA
 ATAATTTGTATGGGATTTGTGATGCAGGAAAGCCTAAGGGAAAAAGAATATTCATTCTGTGTGGT
 GAAAATTTTTTGAAGAAAAAATTGCCTTCTTCAAACAAGGGTGTCAATCTGATATTTATGAGGAC
 TGTGTTCTCACTATGAAGGCATCTGTTATTGAAATGTCTTGTGTTTGTGTTGACTGGAGTAC
 ATTCAAACAAAGAAACGGCAAAGAGATTAAAGGCCCAAGTTCAGTGTGCCTCAGATCAACTGC
 GATGTCAAAGCCGGAAGATCATCGATCCTGAGTTCATTGTGAAATGTCCAGCAGGATGCCAAGA
 CCCCAAATACCATGTTTATGGCACTGACGTGTATGCATCCTACTCCAGTGTGTGTGGCGCTGCCG
 TACACAGTGGTGTGCTTGATAATTAGGAGGGAAAAATCTTGTTCGGAAGGTTGCTGGACAGTCT
 GGTTACAAAGGGAGTTATTTCAACGGTGTCCAATCGTTATCCCTACCACGATGGAGAGAATCCTT
 TATCGTCTTAGAAAGTAAACCCAAAAAGGGTGTAACTACCCATCAGCTCTTACATATTCATCAT
 CGAAAAGTCCAGCTGCCCAAGCAGGTGAGACCACAAAAGCCTATCAGAGGCCACCTATTCCAGGG
 ACAACTGCACAGCCGGTCACTCTGATGCAGCTTCTGGCTGTCACTGTAGCTGTGGCCACCCCCAC
 CACCTTGCCAAGGCCATCCCTTCTGCTGCTTCTACCACCAGCATCCCAGACCACAATCAGTGG
 GCCACAGGAGCCAGGAGATGGATCTCTGGTCCACTGCCACCTACACAAGCAGCCAAAACAGGCC
 AGAGCTGATCCAGGTATCCAAAGGCAAGATCCTTCAGGAGCTGCCTTCCAGAAACCTGTTGGAGC
 GGATGTGACGCTGGGACTTGTTCAAAAGAAGAAATTGAGCACACAGTCTTGGAGCCAGTATCCC
 TGGGAGATCCAACTGCAAAATTGACTTGTGCTTTTTTAATTGATGGGAGCACCAGCATTTGGCAAA
 CGGCGATTCCGAATCCAGAAGCAGCTCCTGGCTGATGTTGCCCAAGCTCTTGACATTGGCCCTGC
 CGGTCCACTGATGGGTGTTGTCCAGTATGGAGACAACCTGCTACTCACTTTAACCTCAAGACAC
 ACACGAATTCTCGAGATCTGAAGACAGCCATAGAGAAAAATTACTCAGAGAGGAGGACTTTCTAAT
 GTAGTTCGGGCCATCTCCTTTGTGACCAAGAACTTCTTTTCAAAGCCAATGGAACAGAAGCGG
 GGCTCCCAATGTGGTGGTGGTGGTGGTGGTGGCTGGCCACGGACAAGTGGAGGAGGCTTCAA
 GACTTGGGAGAGAGTCAGGAATCAACATTTTCTTCACTACCATTAAGGTGCTGCTGAAATGAG
 AAGCAGTATGTGGTGGAGCCCACTTTGCAACAAGGCCGTGTGCAGAACAAACGGCTTCTACTC
 GCTCCACGTGCAGAGCTGGTTTGGCCTCCACAAGACCCTGCAGCCTCTGGTGAAGCGGGTCTGCG
 ACACTGACCGCTGGCCTGCAGCAAGACCTGCTTGAACCTCGGCTGACATTGGCTTCGTCATCGAC
 GGCTCCAGCAGTGTGGGGACGGGCAACTTCCGACCGTCCCTCAGTTTGTGACCAACCTCACC
 AGAGTTTGAGATTTCCGACACGGACACGCGCATCGGGGCCGTGCAGTACACCTACGAACAGCGGC
 TGGAGTTTGGGTTTCGACAAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGGTGGGC
 TACTGGAGTGGTGGCACCAGCAGGGGGCTGCCATCAACTTCGCCCTGGAGCAGCTCTTCAAGAA
 GTCCAAGCCCCAACAAAGAGGAAGTTAATGATCCTCATCACCAGCGGGAGGTCTACGACGACGTCC
 GGATCCAGCCATGGCTGCCATCTGAAGGGAGTGATCACCTATGCGATAGGCGTTGCTGGGCT
 GCCCAAGAGGAGCTAGAAGTCATTGCCACTACCCCGCCAGAGACCCTCTTCTTTGTGGACGA
 GTTTGACAACCTCCATCAGTATGTCCCAGGATCATCCAGAACATTTGTACAGAGTTCAACTCAC
 AGCCTCGGAACGTAATTAGAGCAGGCAGAGCACCAGCAAGTGCTGCTTTACTAACTGACGTGTT
 GGACCACCCACCGCTTAATGGGGCAGCAGCGTGCATCAAGTCTTGGGCAGGGCATGGAGAAAC
 AAATGTCTTGTATTATTCTTTGCCATCATGCTTTTTCATATTCCAAAACCTGGAGTTACAAAGA
 TGATCACAAACGTATAGAATGAGCCAAAAGGCTACATCATGTTGAGGGTGTGGAGATTTACAT
 TTTGACAATTGTTTTCAAATAAATGTTCCGAATACAGTGCAGCCCTTACGACAGGCTTACGTAG
 AGCTTTTGTGAGATTTTAAGTTGTTATTTCTGATTTGAACCTCTGTAACCTCAGCAAGTTTCAT
 TTTTGTATGACAATGTAGGAATTGCTGAATTAATGTTTAGAAGGATGAAAAATAAAAAA
 AA
 AAAG

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR
ESFIVLESKPKKGVITYPSALTYSSSKSPAQAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA
TPTTLPRPSPSAASTTSIPRPQSVGHRSEQEMDLWSTATYTTSSQNRPRADPGIQRQDPGSAAFQKP
VGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATHFNLKHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTLPVKRVCDTDLACSKTCLNSADIGFVIDGSSSVGTGNFRTVLQFVTN
LTKEFEISDTRIGAVQYTYEQRLFEGFDKYSSKPDILNAIKRVGYWSGGTSTGAAINFALEQL
FKKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKGVITYAIGVAWAAQEELEVIATHPARDHSFF
VDEFDNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGTCTTTTTTCTGTTTTTCAGAGTAGTTCACAACAGATCTGAGTGTTTTAATTAAGCATGGAAT
ACAGAAAACAACAAAAAAGCTTAAGCTTTAATTTTCATCTGGAATTCACAGTTTTCTTAGCTCCCTGGACCC
GGTTGACCTGTTGGCTCTTCCCGCTGGCTGCTCTATCACGTGGTGGCTCTCCGACTACTCAGGAGTGTGTA
AAGAACCTTCGGCTCGCGTGCTTCTGAGCTGCTGTGGAATGGCCTCGGCTCTCTGGACTGTCCCTCCGAGTA
GGATGTCACTGAGATCCCTCAAATGGAGCCTCCTGCTGCTGCTCACTCCTGAGTTTCTTTGTGATGTGGTAC
CTCAGCCTTCCCCACTACAATGTGATAGAACGCGTGAAGTGGATGTACTTCTATGAGTATGAGCCGATTTA
CAGACAAGACTTTCACCTTCACACTTCGAGAGCATTCAAAGTCTCTCATCAAAATCCATTTCCTGGTCATT
TGGTGACCTCCCACCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGGTGAAAAAAGTCT
TGGTGGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAGAGGCTGAAAAGGAAGACAAAATGTTGGC
ATTGTCCTTAGAGGATGAACACCTTCTTTATGGTGACATAATCCGACAAGATTTTTTAGACACATATAATA
ACCTGACCTTGAAACCATTTATGGCATTGAGGTGGGTAAGTGTGAGTTTGGCCCAATGCCAAGTACGTAATG
AAGACAGACACTGATGTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGA
GAAGTTTTTTCACAGGTTATCCTCTAATTGATAATTATCCTATAGAGGATTTTACCAAAAACCCATATTT
CTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTATATAATGTCCAGAGAT
TTGGTGCCAAGGATCTATGAAATGATGGGTACGTAAGAACCCATCAAGTTTGAAGATGTTTATGTCCGGAT
CTGTTTGAATTTATTAAAAGTGAACATTCATATTCAGAAGACACAAATCTTTCTTCTATATAGAATCC
ATTTGGATGTCTGTCAACTGAGACGTGTGATTGCAGCCATGGCTTTTCTTCCAAGGAGATCATCACTTTT
TGGCAGGTCATGCTAAGGAACACCACATGCCATTATTAAGTTTCACTTACATTCTACAAAAGCCTAGAAGGACAG
GATACCTTGTGAAAGTGTAAATAAAGTAGGTACTGTGGAATTCATGGGGAGGTCAGTGTGCTGGCTT
ACACTGAACTGAACTCATGAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTCAGG
CCCTTCAAAGATGATATGTGGAGGAATTAATATAAAGGAATTGGAGGTTTTTGCTAAAGAAATTAATAGG
ACCAAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAGGGTGTACTGAGTTATAAGCTCA
CTAGGCTGTAAAACAAAACAATGTAGAGTTTATTATTGAACAATGTAGTCACTTGAAGGTTTTGTGTA
TATCTTATGTGGATTACCAATTTAAAAATATATGTAGTTCTGTGTCAAAAAACTTCTTCACTGAAGTTATA
CTGAACAAAATTTACCTGTTTTTGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTACAGTTATT
ATTATTTAAATTAAGTTCACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAGGATAG
TGAATCATTCTTTACATGCAACATTTTCCAGTTACTTAACTGATCAGTTTATTATTGATACATCACTCCA
TTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGGACTTTGTAAATATTTTACTGTGGT
AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPiYRQDFHF
TLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQAEKEDKMLA
LSLEDEHLLYGDIIRQDFLDTYNNLTTLKTIMAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYL
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRDLVPRIYEMMGHV
KPIKFEDVYVGICLNLKVNIPEDTNLFFLYRIHL DVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCAGCCGCGCAAGGATGGAGCTGGGTTGCTGGACGCAGTTGGGGCTCACTTTTCTTCAGCTCCTTCTCATC
TCGTCCTTGCCCAAGAGAGTACACAGTCATTATGAAGCCTGCCCTGGAGCAGAGTGGAATATCATGTGTCGGGAGTGCTGTG
AATATGATCAGATTGAGTGCCTCTGCCCCGAAAGAGGGAAGTCGTGGGTTATACCATCCCTTGCTGCAGGAATGAGGAGAA
TGAGTGTGACTCCTGCCTGATCCACCCAGGTTGTACCATCTTTGAAAAGTGCAGAGCTGCCGAAATGGCTCATGGGGGGT
ACCTTGGATGACTTCTATGTGAAGGGTTCTACTGTGCAGAGTGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTG
GCCAGGTTCTGCGAGCCCCAAGGGTCAGATTTTGTGGAAGCTATCCCTTAAATGCTCACTGTGAATGGACCATTTCATGC
TAAACCTGGGTTTGTCTCACTCAACTAAGATTTGTCTATGTTGAGTCTGGAGTTTGACTACATGTGCCAGTATGACTATGTTGAG
GTTCTGATGGAGACAACCGCATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAG
GATCCTCACTCCACGTCCTCTTCCACTCCGATGGCTCCAAGAATTTTGACGGTTTCCATGCCATTTATGAGGAGATCACAGC
ATGCTCCTCATCCCTTGTTCATGACGGCACGTGCGTCCCTTGACAAGGCTGGATCTTACAAGTGTGCTGCTTGGCAGGC
TATACTGGGCAGCGCTGTGAAAATCTCCTTGAAAGAAAGAACTGCTCAGACCTGGGGGCCAGTCAATGGGTACCAGAAAA
TAACAGGGGGCCCTGGGCTTATCAACGGACGCCATGCTAAAAATTGGCACCGTGGTGTCTTTCTTTTGTAACTCCTATGT
TCTTAGTGGCAATGAGAAAAGAACTTGCCAGCAGAAATGGAGAGTGGTCAGGGAAACAGCCCATCTGCATAAAAGCTGCCGA
GAACCAAGATTTCAGACCTGGTGAGAAAGGAGTTCTTCCGATGCAGGTTTCAAGGGAGACACCATTACACAGCTAT
ACTCAGCGGCCTTCAGCAAGCAGAACTGCAGAGTGCCTTACCAAGAAGCCAGCCCTTCCCTTTGGAGATCTGCCATGGG
ATACCAACATCTGCATACCCAGCTCCAGTATGAGTGCATCTCACCCTTCTACCGCCGCTGGGCAGCAGCAGGAGGACATGT
CTGAGGACTGGGAAGTGGAGTGGCGGGGCACCATCTGCATCCCTATCTGCGGGAAAAATTGAGAATCACTGCTCCAAAGA
CCCAAGGGTTGCGCTGGCGTGGCAGGCAGCCATCTACAGGAGGACCAGCGGGGTGCATGACGGCAGCCTACACAAGGGAGC
GTGGTTCTTAGTCTGAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCACTGTGTTACTGACCTGGGGAAG
GTCACCATGATCAAGACAGCAGACCTGAAAGTTGTTTGGGGAAATCTACCAGGATGATGACCGGGATGAGAAGACCATCC
AGAGCCTACAGATTTCTGTATCATTCGATCCCAACTATGACCCCATCCTGCTTGATGCTGACATCGCCATCCTGAAGCT
CCTAGACAAGGCCCTGATCAGACCCGAGTCCAGCCCATCTGCTCGCTGCCAGTGGGATCTCAGCACTTCTTCCAGGAG
TCCCACATCACTGTGGCTGGCTGGAATGCTTGGCAGAGCTGAGGAGCCCTGGCTTCAAGAACGACACACTGCGCTCTGGGG
TGGTCAGTGTGGTGGACTCGCTGCTGTGTGAGGAGCAGCATGAGGACCATGGCATCCAGTGAGTGTCACTGATAACATGTT
CTGTGCCAGCTGGGAACCCACTGCCCTTCTGATATCTGCATGACAGAGACAGGAGGCATCGCGGCTGTCTCTTCCCGGA
CGAGCATCTCCTGAGCCACGCTGGCATCTGATGGGACTGGTCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCA
CTGCCCTTACCAAGGTGCTGCCCTTTTAAAGACTGGATTGAAAGAAATATGAAATGAACCATGCTCATGCACTCCTTGAGAAG
TGTTTCTGTATATCCGTCTGTACGTGTGTCATTGCGTGAAGCAGTGTGGGCTGAAGTGTGATTGGCCTGTGAACCTGGCT
GTGCCAGGGCTTCTGACTTCAGGGACAAAAGTCAAGTGAAGGGTGAGTAGACCTCCATTGCTGGTAGGCTGATGCCGCGTCCA
CTACTAGGACAGCCAATTGGAAGATGCCAGGGCTTGCAAGAAGTAAGTTCTTCAAGAAGACCATATACAAAACCTCTCCA
CTCCACTGACCTGGTGGCTTCCCCAATTTTCAAGTTATACGAATGCCATCAGCTTGACCAGGGAAGATCTGGGCTTCATGAG
GCCCCCTTTGAGGCTCTCAAGTTCTAGAGAGTGCCTGTGGGACAGCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCCCTT
TGTGTACATGGCCACAGTACAGTCTGCTCTTTTCTTCCCCATCTTGTACACATTTTAAATAAATAAGGGTTGGCTTCT
GAACTACAAAAA
AAAAA

FIGURE 38

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVVGYT
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLDDFYVKGIFYCAECRAGWYGGDCMRCGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGQII
KRVCGNERPAPIQSIGSSSLHVLFSHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE
KRTCQONGEWSGKQPICIKACREPKISDLVRRRVLPQVQSRETPLHQLYSAAFSKQKLQSAPTK
KPALPFGDLPMDGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGKWSGRAPSCIPICGKIENITAP
KTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAHCVTDLGKVTMIKTADL
KVVLGKFYRDDRDEKTIQSLQISAILHPNYDPILLDADIAILKLLDKARISTRVQPICLAASR
DLSTSFQESHITVAGWNVLADVRSPGFKNDTLRSGVSVVDSLLCEEQHEDHGIPVSVTDNMFCA
SWEPTAPSDICTAETGGIAAVSFPGRASPEPRWHLMLVSWSYDKTCSHRLSTAFTKVLFPKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTTCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCCGTGATTTATTAACGTGGCTTAATC
TGAAGGTTCTCAGTCAAATTTCTTTGTGATCTACTGATTGTGGGGGCATGGCAAGGTTTGCTTAAAGGAGCTTGGCTGG
TTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAATGAAGGCGCTTCTGTTGC
TGGTCTTGCCTTGGCTCAGTCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCTCTGTATTAGAACTCTGTA
AAGGTGCCCTCCACTACGGCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTGTCCAGACGGCTGTGCGAGCC
TCACAGCCACGGCTCCCTCCCAAGGTTTCTGCAGCTGCCACCATCTCTTAATGACAGACGAGCTGGCCTAGACA
ACCCCTGCCTACGTGTCTCGGCAGAGGACGGGCAGCCAGCAATCAGCCAGTGGACTCTGGCCGGAGCAACCGAACTA
GGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAAAAAATAAATCGAGCTTTGAGTGTTCCTCGAA
GGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGGGCAGGGAATTTCTGAAAACACCACTGCCCTGAAG
TCTTTCCAAGGTTGTACCACCTGATTCCAGATGGTGAATTAACAGCATCAAGATCAATCGAGTAGATCCAGTGAAA
GCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCACTGGTCCATATCATTATCCAACACATTTATCTGATGGGG
TGATCGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTC
ACAACTAGCTGTGCGTCTCTGCGGCAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAGTCCGCA
GCAGGAACAATGGACAGGCCCCGGATGCCACAGCCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCC
CCGAGGAGCAGCTTGGAAATAAACTGGTGCAGAGGTGGATGAGCCTGGGGTTTTTCATCTTCAATGTGCTGGATGGCG
GTGTGGCATATCGACATGGTCACTTGGAGAGAATGACCGTGTGTAGCCATCAATGGACATGATCTTCGATATGGCA
GCCCAGAAAGTGCGGCTCATCTGATTCAAGCCAGTGAAAGACGTGTTCACTCGTGTGTCGCCGAGGTTCCGCGAGC
GGAGCCCTGACATCTTTAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCCAGGGGAGAGGAGCAACA
CTCCCAAGCCCTCCATCCTACAATTACTTGTCTATGAGAAGGTGGTAAATATCCAAAAAGACCCCGGTGAATCTCTCG
GCATGACCGTCGAGGGGAGCATCACATAGAGAATGGGATTTGCCTATCTATGTCATCAGTGTGAGCCCGGAGGAG
TCATAAGCAGAGATGGAAGAATAAAACAGGTGACATTTTGTGAATGTGGATGGGGTCGAACAGAGGTGAGCC
GGAGTGAGGCAGTGGCATTATTGAAAAGAACATCATCTCGATAGTACTCAAAGCTTTGGAAGTCAAAGAGTATGAGC
CCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCAGTGACTGGTCCCCATCCT
GGGTGATGTGGCTGGAATTACCACGGTGTGTATAACTGTAAAGATATTGTATTACGAAGAAACAGCTGGAAGTC
TGGGCTTCTGCATTGTAGGAGGTTATGAAGAATACAATGGAACAAACCTTTTTTCATCAAATCCATTGTTGAAGGAA
CACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTGCTGTCAATGGTAGAAGTACATCAGGAATGA
TACATGCTTGCTTGGCAAGACTGCTGAAAGAACTTAAAGGAAGAATTACTCTAACTATTGTTTCTTGGCCTGGCACTT
TTTTATAGAATCAATGATGGGTGAGGAGAAACAGAAAAATCACAAATAGGCTAAGAAGTTGAAACACTATATTTATC
TTGTGATTTTTATATTTAAAGAAAGAAATACATTGTAAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTT
ACACCTCAGAAAATATGATTCAAAAAAATTAATACTACTAGTTTTTTTTTCAGTGTGGAGGATTTCTATTACTCTAC
AACATTGTTTATATTTTTCTATTCAATAAAAAGCCCTAAACAACATAAATGATTGATTGTATACCCCACTGAATT
CAAGCTGATTTAAATTTAAATTTGGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTTTAAATTTACAGCT
AAAAATTTTTTAAATGCAATGCTGAGAAACGTTGCTTTCATCAAACAAGAATAAATATTTTTCAGAAGTTAAA

FIGURE 40

MKALLLLVLPWLSpanyIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPCDGCASLTATAPS
PEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRSFKKINR
ALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV
GGSETPLVHIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGVFIFNVLDGGVAYRHG
QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQVRQSPDIFQEAGWNSNGSWSPG
PGERSNTPKPLHPTITCHEKVNIQKDPGESLGMTVAGGASHREWDLPYVISVEPGGVISRDGR
IKTGDIILLNVDGVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP
SDWSPSWMMWLELPRCLYNCKDIVLRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
RIRCGDILLAVNGRSTSGMIHACIARLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTT
CTTGCCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAACATGGGCTTCAACCTGACT
TTCCACCTTTCTACAAATCCGATTACTGTTGCTGTTGACTTTGTGCCTGACAGTGGTTGGGTGGGC
CACCAGTAATACTTCTGTTGGTGGCATTCAAGAGATTCTTAAAGCAAAGGAGTTTCATGGCTAATTTCC
ATAAGACCCTCATTTTGGGGAAGGGAAAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAACTT
GACAACTGTCTTCTGTGTCTCCTTACCTCAGAGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCAC
TTTGAAGAGGTACAGGCAGAAAATCCCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAG
CTTTACAGAGGGTCGCCATCCTCGTTCCCCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAA
CATCTGCATCCCTTCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGG
TAAAAAGTTTAAATCGAGCCAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGG
ACTGCTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGAGGAG
CATCCCAAGCATCTGGTGGTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTGGATATTTTGG
GGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCTAACAATACTGGGGAT
GGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTTCCCGGCCCTG
CCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACG
GATGAAGCTCTTACACCAAGTGTACGAGTCTGGAGAACAGATGGGTTGAGTAGTTGTTCTTATAAAT
TAGTATCTGTGGAACACAATCCTTTATATATCAACATCACAGTGGATTTCTGGTTTGGTGCATGACCC
TGGATCTTTTGGTGATGTTTGAAGAAGTATTCTTTGTTTGAATAATTTTGGCCTAGAGACTTCAA
ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTTCCTTTTGTATTTCT
TAGCAGAGCTCCTGGTGATGTAGAGTATAAACAGTTGTAACAAGACAGCTTCTTAGTCATTTTGAT
CATGAGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAGGATGACTCAAAGGATAA
AATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATATTATGGGAT
AAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGTAGAA
AGGTACGAAGATACAATACTGTTATTTCATTTATCCTGTACAATCATCTGTGAAGTGGTGGTGTCAAGT
GAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAAGTTGGGAATGAAGA
GGTAGCAGGAGGGTGGAGTGTGGGCTGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC
CTTCAGGGGAGGACCTGCCAGGTATGCCTTCCAGTGATGCCACCAGAGAATACATTCTCTATTAGT
TTTTAAAGAGTTTTTGTAAAATGATTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACAT
ATTAATAATAATAAATATGTCTATCAAATACCTCTGTAGTAAATGTGAAAAAGCAAAA

FIGURE 42

MGFNLTfHLSYKfRLLLLLTCLTVVGWATSNYFVGAIQEIpkakefMANfHKTLILGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEgkKfNRAKLLNVGYLEALKEENWDCFI FHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTALSREQFFKVNfGSNNYWGwGGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGCACTGAGCTCCCAGATCTGGG
CCGCTTGCCTCCTGCTCCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCACAACAG
ACGGGACAACTTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCAT
GTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTTCTGCTGCGGCTGCTGTC
ATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCCGTCCCCTCCC
TTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGGATAAAATGGCTGGTTCTTTGTTT
TCCAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 44

MALSSQIWAACLLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFPI
CIFCCGCCHRSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCACCC
TCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT
TCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTATTGTCTG
GACCTTCAACACAACCCCTCTTGTCAACATACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACCTGAAG
AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC
CTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTCAGCAGAACT
TCTCAAGCCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATG
GTCTCTCTGTGTCTCTCTGTTGGTGGCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTGGG
AAACTCCTAACATATGCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGAT
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAAATGTTA
TCTAGACAGCAGTGCCTCCCCTAAGTCTCTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVTIQP
EGGTIIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCMEHGEEVDIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLWFLKRERQEEYIE
EKKRVDICRETPNICPHSGENTYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTTCTCAAAACCCCATCTCTTGCTTTGAGTGGTGGTTCCAGGAATT
ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTCATCCAGAATCCTTCAA
CTTGCACTGGTTTTTCAATGACTCTTGTCACCTCCTACTGGTTTCAATAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTCCACTTCGATTCTGAAGAAACAAACATAGG
CTTATCCACTTCTCAGTATTTTTAGGTCTATTGCTTGTTGGAATTCTGGAGGTCTGTTTGGGCT
CAGTCAGATAGTCATCGGTTTCCTTGGCTGTCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG
TGTAGTTTTAATGGGAATAAAATGTAAGTATCAGTAGTTTGAAAAAAAAA

FIGURE 48

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCLISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSSEENKHRLIHFSVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTCGAACTGTGA
CATGGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGGAAGCCAATGACC
CATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTCTGCCA
CTACTTGCT**AG**GCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACC
TCCTCCCCCTGGGAGGCCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCTT
TTCTGATCAGGAGGCTTCTTTATGAATTAAACTCGCCCCACCACCCCTCA

FIGURE 50

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAAVLSGKCKYKS
SQKQHSPVPEKAIPITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG
AGGAGACCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGGCAGAGATGAAGTTCCAG
GGGCCCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGGCAGTGGGAGGCTGGCCCCCTGCAGAG
CGGAGAGGAAAGCACTGGGACAAATATTGGGGAGGCCCTTGACATGGCCTGGGAGACGCCCTGA
GCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGGAGGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTTGGCGC
AGCAGATGCTTTGGGCAACAGGGTCGGGGAAGCAGCCCATGCTCTGGGAAACACTGGGCACGAGA
TTGGCAGACAGGCAGAAGATGTCATTGACACGGAGCAGATGCTGTCCGCGGCTCCTGGCAGGGG
GTGCCTGGCCACAGTGGTGCTTGGGAACTTCTGGAGGCCATGGCATCTTTGGCTCTCAAGGTGG
CCTTGAGGGCCAGGGCCAGGGCAATCCTGGAGGTCTGGGGACTCCGTGGGTCCACGGATACCCCG
GAAACTCAGCAGGCAGCTTTGGAATGAATCCTCAGGGAGCTCCCTGGGGTCAAGGAGGCAATGGA
GGGCCACCAAACCTTTGGGACCAACACTCAGGGAGCTGTGGCCAGCCTGGCTATGGTTTCAGTGAG
AGCCAGCAACCAGAATGAAGGGTGCACGAATCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGGAGGCAGCGGCTCACAGTCGGGCAGCAGTGGCAGTGGCAGCAATGGTGACAACAAC
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAG
TGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGT
CCTCCTGGGGATCCAGCACCGGCTCCTCCTCCGCAACCACGGTGGGAGCGGCGGAGGAAATGGA
CATAAACC CGGTGTGAAAAGCCAGGGAATGAAGCCCGCGGGAGCGGGGAATCTGGGATTCAGGG
CTTCAGAGGACAGGGAGTTTCCAGCAACATGAGGGAATAAGCAAAGAGGGCAATCGCCTCCTTG
GAGGCTCTGGAGACAATTATCGGGGGCAAGGGTCGAGCTGGGGCAGTGGAGGAGGTGACGCTGTT
GGTGGAGTCAATACTGTGAACCTCTGAGACGTCTCCTGGGATGTTTAACTTTGACACTTTCTGGAA
GAATTTTAAATCCAAGCTGGGTTTCATCAACTGGGATGCCATAAACAAGGACCAGAGAAGCTCTC
GCATCCCGTGACCTCCAGACAAGGAGCCACCAGATTGGATGGGAGCCCCCAGCTCCCTCCTTAA
AACACCACCCTCTCATCACTAATCTCAGCCCTTGCCCTTGAAATAAACCTTAGCTGCCCCACAAA
AA
AA

[illegible]

Signal peptide:

N-glycosylation site.

Glycosaminoglycan attachment site.

Casein kinase II phosphorylation site.

N-myristoylation site.

Cell attachment sequence.

APP ID=10063599

FIGURE 53

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCCTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCT
ACTCGCCCGCATCCTGGCTTGACCTATGCCTTCTATAACAACCTGCCGCCGGCTCCAGTGTTTCC
CACAGCCCCAAAACGGAACGGTTTTGGGGTCACCTGGGCCTGATCACTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTGGCCACCTATTCCCAGGGCTTTACGGTATGGCTGGGTCCCAT
CATCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCACCAATGCCTCAGCTGCCA
TTGCACCCAAGGATAATCTCTTCATCAGGTTCTGAAGCCCTGGCTGGGAGAAGGGATACTGCTG
AGTGGCGGTGACAAGTGGAGCCGCCACCGTCCGATGCTGACGCCCCGCTTCCATTTCAACATCCT
GAAGTCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAGCACATCAGCCTCATGACCTTGGACAGTCTA
CAGAAATGCATCTTCAGCTTTGACAGCCATTGTCAGGAGAGGCCCAGTGAATATATTGCCACCAT
CTTGGAGCTCAGTGCCCTTGTAGAGAAAAGAAGCCAGCATATCCTCCAGCACATGGACTTTCTGT
ATTACCTCTCCCATGACGGGCGGCGCTTCCACAGGGCCTGCCGCTGGTGCATGACTTCACAGAC
GCTGTCATCCGGGAGCGGCGTCGCACCCTCCCCACTCAGGGTATTGATGATTTTTTCAAAGACAA
AGCCAAGTCCAAGACTTTGGATTTCAATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGG
CATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGAGAGCCATGACACCACG
GCCAGTGGCCTCTCCTGGGTCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG
ACAGGAGGTGCAAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCC
AGCTGCCCTTCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCAGCTCCCTTCATC
TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTG
CCTCATCGATATTATAGGGGTCCATCACAACCCAACTGTGTGGCCGGATCCTGAGGTCTACGACC
CCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTACCTCTGGCTTTTATTCTTTCTCCGCA
GGGCCCAGGAACATGCATCGGGCAGGCGTTGCCATGGCGGAGATGAAAGTGGTCCTGGCGTTGAT
GCTGCTGCACTTCCGGTTCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCTGAATGTAGGCTTGCACTGACTTTCTGAC
CCATCCACCTGTTTTTTTGCAGATTGTCATGAATAAAACGGTGCTGTCAA

FIGURE 54

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPPQPPKRNWFWGHLG
LITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
WLGE GILLSGGDKWSRHRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRAC
RLVHDFTDVIRERRRTLPTQGIDDFKDKAKSKTLD FIDVLLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLRHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENSKGRSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTTGA
GCCTACTCGTTTGATTGCAACTATCATGGTGCTGTTGTGTTTTGCACTTACCCTGTGTTCTGCCT
TTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTTGCACTCTTTGGCATTGACGTGG
TACAGCCTTTCCTTCATACCATTTGCAAGGGATGCTGTGAAGAAGTGTGTTGCCGTGTGTCTTGC
ATAATTTCATGGCCAGTTTTATGAAGCTTTGGAAGGCACTATGGACAGAAGCTGGTGGACAGTTTT
GTAAGTATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCTTGACGAATGTGTGCTT
GTGATTCGAACATTTGAGGGTTACTTTTGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
GCACAGGATGAGAAGTGGGTTCTGTATCTTGTGGAGTGGAAATCTTCCTCATGTACCTGTTTCCTC
TCTGGATGTTGTCCCACTGAATTCCCATGAATACAAACCTATTTCAGCAACAGCAAAAAAAAAAAAA
AA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGALIFCILQSLALTWYSLSFIPFAR
DAVKKCFVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCGAGTGGAGCACCCAGCAGGCCGCCAACAATGCTCTGTCTGTGCCTG
TACGTGCCGGTCAATCGGGGAAGCCAGACCGAGTTCAGTACTTTGAGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCC
ATTTTCAAGCTCAGTGTCTTATCCCCCTCCAGGAATTCTCCACCTACCGCCAGTGGAGCAGAAAATTGTACAAGCT
GGAGATAAGGACCTTGATGGGCAGCTAGACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGAGG
CTGGTGTTTAAGATTTTGGACAAAAGAATGATGGACGCATTGACGCGCAGGAGATCATGCAGTCCCTGCGGGACTTG
GGAGTCAAGATATCTGAACAGCAGGCAGAAAAAATTCGAAGAGCATGGATAAAAACGGCAGCATGACCATCGACTGG
AACGAGTGGAGAGACTACCACCTCCTCCACCCCGTGGAAAACATCCCCGAGATCATCCTCTACTGGAAGCATTCACAG
ATCTTTGATGTGGGTGAGAATCTAACGGTCCCGGATGAGTTACAGTGGAGGAGAGGCAGACGGGGATGTGGTGGAGA
CACCTGGTGGCAGGAGTGGGGCAGGGGCCGTATCCAGAACCTGCACGGCCCCCTGGACAGGCTCAAGGTGCTCATG
CAGGTCCATGCCCTCCGCGAGCAACAACATGGGCATCGTTGGTGGCTTCACTCAGATGATTGAGAAGGAGGGGCCAGG
TCACCTCTGGCGGGCAATGGCATCAACGTCTCAAAATGCCCCCGAATCAGCCATCAAATTCATGGCCATGAGCAG
ATCAAGCGCCTTGTGGTAGTGACAGGAGACTCTGAGGATTACGAGAGGCTTGTGGCAGGGTCTTGGCAGGGGCC
ATCGCCAGAGCAGCATCTACCCAAATGGAGGTCTGAAGACCCGGATGGCGCTGCGGAAGACAGGCCAGTACTCAGGA
ATGCTGGACTGCCCCAGGAGGATCCTGGCCAGAGAGGGGTGGCCGCTTCTACAAAGGCTATGTCCCCAACATGCTG
GGCATCATCCCCATGCCGCGCATGCACCTTGCACTACGAGACGCTCAAGAATGCCTGGCTGCAGCACTATGCAGTG
AACAGCGCGGACCCCGCGTGTGGTGTCTGGCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGGCCAGCTAC
CCCCGGCCCTAGTCAGGACCCGGATGCAGGCGCAAGCCTCTATTGAGGGCGCTCCGGAGGTGACCATGAGCAGCCTC
TTCAAAATATCTTCCGAGCCGAGGGGGCCCTTCGGGCTGTACAGGGGGCTGGCCCCAACTTCATGAAGGTCAATCCCA
GCTGTGAGCATCAGCTACGTGGTCTACGAGAACCTGAAGATCACCTGGGCGTGCAGTCGCGGTGACGGGGGAGGGC
CGCCCCGCGAGTGGACTCGCTGATCCTGGGCGCAGCCTGGGGTGTGCAGCCATCTCATTCTGTGAATGTGCCAACT
AAGCTGTCTCGAGCCAGCTGTGAAAACCTAGACGCACCCGAGGGAGGGTGGGGAGAGCTGGCAGGCCCAGGGCTT
GTCTGTCTGACCCAGCAGACCCCTCTGTTGGTTCAGCGAAGACCACAGGCATTCTTAGGGTCCAGGGTCAGCAGG
CTCCGGGCTCACATGTGTAAAGCACAGGACATTTCTGCACTGCCTGCCAATAGTGAGCTTGGAGCCTGGAGGCCGGCT
TAGTCTCTCCATTTCACCCCTGACGCCAGCTGTTGGCCACGGCCCTGCCCTCTGGTCTGCCGTGCATCTCCCTGTGC
CCTCTGTCTGCCTGCTGTCTGCTGAGGTAAGGTGGGAGGAGGGCTACAGCCACATCCACCCCTCTGCTCAATCCC
ATAATCCATGATGAAAGGTGAGGTACGTGGCCTCCAGGCCTGACTTCCCAACCTACAGCATTGACGCCAATTTGGC
TGTGAAGGAAGAGGAAAGGATCTGGCCTTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGGCTCTCGGCATG
CTTGGGAGTGCAGGGGGCTCGGGCTGCCTGGCTGGCTGCACAGAAGGCAAGTGTGGGGCTCATGGTGTCTGAGCT
GGCCTGGACCCTGTGAGGATGGGCCCCACCTCAGAACCAACTCACTGTCCCCACTGTGGCATGAGGGCAGTGGAGCA
CCATGTTTGAAGGCGAAGGGCAGAGCGTTTGTGTCTGGGGAGGGAAAGGAGGTTGGAGGCCCTTAATTATGG
ACTGTTGGGAAAAGGGTTTGTCCAGAAGGACAGCCGACAAATGAGCGACTTCTGTCTTCCAGAGGAAGACGAGG
GAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGGTTCTGTCCAAACCCAGCAGGGGCGCAGC
GGACCAGCCCCACATTCCACTTGTGTCACTGCTTGAACCTATTTATTTTGTATTATTGTAACAGAGTTATGTCTCT
AACTATTTTATAGATTTGTTTAATTAATAGCTTGTCTTTTCAAGTTCATTTTTATTATCATATTTATGTTTATGGTT
GATTGTACCTTCCCAAGCCCCCAGTGGGATGGGAGGAGGAGGAGAGGGGGGCTTGGGGCGCTGCAGTCACATCT
GTCCAGAGAAATTCCTTTTGGGACTGGAGGCAGAAAAGCGGCCAGAAAGGCAGCAGCCCTGGCTCCTTTCTTTGGCAG
GTTGGGGAAGGGCTTGGCCCCAGCCTTAGGATTTTCAAGGTTTGAAGTGGGGGCTGGAGAGAGAGGGAGAACCTCAAT
AACCTTGAAGGTGGAATCCAGTTATTTCTGCGCTGCGAGGGTTTCTTTATTTCACTCTTTTCTGAATGTCAAGGCAG
TGAGGTGCCCTCTCACTGTGAATTTGTGGTGGGCGGGGGCTGGAGGAGAGGGTGGGGGGCTGGCTCCCTCCCTCCAGC
CTTCTGCTGCCCTTGTCTTAAATATGCCGCCAACTGGCGACCTCACGGTTGCACTTCCATTCCACAGAAATGACCTGA
TGAGGAAATCTTCAATAGGATGCAAGATCAATGCAAAAATTTGTATATATGAACATATAACTGGAGTCTGTCAAAAAG
CAAAATTAAGAAAGAATTGGACGTAGAAGTTGTCAATTAAGCAGCCTTCTAATAAAGTTGTTTCAAGCTGAAAAAA
AA

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSI FKLSVFIPSQEFSTYRQWKQKIVQAGDKDLG
QLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILKSMDKNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTPDEFTVEERQTGMWWRHLVAGGG
AGAVSRTCTAPLDRLKVLMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIAK
FMAYEQIKRLVGSQDETLEIRLHVAGSLAGAIQAQSSIYPMEVLKTRMALRKTGQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNALQHYAVNSADPGVFVLLACGTMSSTC
GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCATGGC
TTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTTGGTATTTTCAGGGAGACACTCCATCACAGTCACTACTGTGCGCTCAGCT
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTTGAACCTGACATCAAACCTTTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTGCGAGCAGGATGAAATGTTTCAGAGGCCGGACAGCAGTGTGCTGATCAAGTGATAGTT
GGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAACTGGAGCCTTCAGCATGCCGG
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCAGGTGTGAGGCTCCCCGATGGTTCCCC
CAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCAGGGAGCCAACCTTCTCGGAAGTCTCCAATAC
CAGCTTTGAGCTGAACTCTGAGAATGTGACCATGAAGGTTGTGTCTGTGCTCTACAATGTTACGA
TCAACAACACATACTCCTGTATGATTGAAAAATGACATTGCCAAAGCAACAGGGGATATCAAAGTG
ACAGAATCGGAGATCAAAGGCGGAGTCACCTACAGCTGCTAAACTCAAAGGCTTCTCTGTGTGT
CTCTTCTTTCTTTGCCATCAGCTGGGCACCTTCTGCCTCTCAGCCCTTACCTGATGCTAAAATAAT
GTGCCTTGGCCACAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTTAC
CACCAGATATGACCTAGTTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAACAAAAAGAAGCCAAAGCAGAAGGCTCCAATATGAACAAGATAAAT
CTATCTTCAAAGACATATTAGAAGTTGGGAAAATAATTTCATGTGAAGTAGACAAGTGTGTTAAGA
GTGATAAGTAAAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT
CACCTGGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTTAGTTATATGTGC
TGTAATGTTGCTCTGAGGAAGCCCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATCCAC
AAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGGGGCG
GCTGCATTTTAGTAATGGGTCAAATGATTCACTTTTTATGATGCTTCCAAAGGTGCCTTGGCTTC
TCTTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTAGCATAAACAGAGCAGT
CGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTAAACAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 60

MASLGQILFWSIIISIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAGTYKC
YIITSKKGKNANLEYKTGAFSMPEVNVVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVVSVLNVNTINNTYSCMIENDIAKATGDIKVTESI KRRSHLQLLSKASL
CVSSFFAISWALLPLSPYLMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCCAGGAGCTGCAGGACAAG
CACCAGGAGCCCCCTCCGGGTAGCTACTACCCTGGACCCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCCTGGTGGTGGTTATGGGGGTCTGCCCTGGAGGGCCTTATGGACCACCAGCTGG
TGGAGGGCCCTATGGACACCCCAATCCTGGGATGTTCCCCTCTGGAACCTCAGGAGGACCATATG
GCGGTGCAGCTCCCGGGGGCCCCCTATGGTCAGCCACCTCCAAGTTCCTACGGTGCCAGCAGCCT
GGGCTTTATGGACAGGGTGGCGCCCCCTCCCAATGTGGATCCTGAGGCCTACTCCTGGTTCCAGTC
GGTGGACTCAGATCAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGCTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTTGACAAGACCAAGTCA
GGCCGCATCGATGTCTACGGCTTCTCAGCCCTGTGGAAATTCTCCAGCAGTGGAAGAACCTCTT
CCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC
AAATGGGCTACAACTGAGCCCCCAGTTCACCCAGCTTCTGGTCTCCCGCTACTGCCACGCTCT
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
GGCCTTCCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTTCGAGGACTTCGTCA
CCATGACAGCTTCTCGGATGCTATTCACCCAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT
TCCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTCTTTTCTGTCCCTCTAGAAGAAC
ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAAGAGGGTGGAGAGTCTTGCATCATAGCCACCA
AATAGTGAGGACCGGGGCTGAGGCCACACAGATAGGGGCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCTTGTAAATGG
AGTTAGTGTCAGTCAGCTGAGCTCCACCCTGATGCCAGTGGTGAGTGTTCATCGGCCTGTTACC
GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAACGAGCCCATTTTCTCCAAAGTGAAT
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAGTGGCTTGGATTCTGCCACACCCATAAAT
CCTTGTTGTGTTAACTTCTAGCTGCCTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCAT
CTTTGGCCAGGCTTCTGCCCTGCAGCTGGGACCCCTCACTTGCTGCCATGCTCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTTAAATTGCATTTTTTTC
ATTTGGGGCCAAAAGTCCAGTGAAATTGTAAGCTTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSGTPGGPYGGAAPGGPYGQPPSSYGAQQPGLYGQGGAPPNVDP EAYSWFQSVDSH
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSALWKFIQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC
TCCCAGGCCTCTTTGCCCGGAGCATCGGTGTTGTGGAGGAGAAAGTTTCCCAAACTTCGGGACC
AACTTGCTCAGCTCGGACAACCTTCCTCCACTGGCCCTCTAACTCTGAACATCCGCAGCCCGC
TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTTCCCTCTGAAGCTCAGCGTGCCTCCATCAGATG
GCTTCCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGGCTGCCTGCCATG
GATTCTTGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCCTGGGGGA
AGCGCTGCCTGAAGAACTCTCTTACCTCTCCAGTGTGCGGCCCTCGCTCCGGGCAGTGGCCCTT
TGCCTGGGGAGTCTTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCTCCACCAGGAC
TCGGAGTCCAGACGACTGCCCCGTCTAATTCACTGGGAGCCGGGGGAAAAATCCTTTCCAACG
CCCTCCCTGGTCTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGGTACCCTGAATCCAGTG
TGTCTGGGGAGGTGGAGGCCCTGGGACTGGTGGGGAACGAGGCCCATGCCACACCCTGAGGGA
ATCTGGGGTATCAATAATCAACCCCCAGGTACCAGCTGGGGAAATATTAATCGGTATCCAGGAGG
CAGCTGGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGGGAATATTAATCGGTATCCAGGAG
GCAGCTGGGGGAATATTATCTATACCCAGGTATCAATAACCCATTTCCTCCTGGAGTTCTCCGC
CCTCCTGGCTCTTCTTGGAACATCCCAGCTGGCTTCCCTAATCCTCCAAGCCCTAGGTTGCAGTG
GGGCTAGAGCACGATAGAGGGAAACCAACATTGGGAGTTAGAGTCTGCTCCCGCCCTTGCTG
TGTGGGCTCAATCCAGGCCCTGTTAATGTTTCCAGCACTATCCCCACTTTTCAGTGCCTCCCC
TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 64

MQGRVAGSCAPLGILLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGSPNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPPAGGSAVQRWPPSWGLPAMDSWPPEDPWQMAAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRNSNLGAGGKILSRP
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRPMPHPEGIWGINNQPPGTSGNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPPGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGGTC
TGGGCTGCCCCCTTGTCCTCCTCTTGACCTCCTTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGCTTCCTGGAA
TTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCACCCTCCACCATGC
AAGATCTCAACACCATGTTGTCTGCAACACATTGACAGCCATTGAAGCCTGTGTCTTCTTGGCCC
GGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCGACCCTGTCTTTCAGCAGGCCCCCACCCTC
CTGAGTGGCAATAAATAAAATTCGGTATGCTG

FIGURE 66

MGSGLPVLVLLLTLLGSSSHGTGPGMTLQLKLKESFLTNSSYESSFLELLEKLCLLLHLPSTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCGGGC
CAGGTGCCCCGTGCGAGGTGCCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGAAGCCCC
TTCTCGGCGCTGCCAACCCGCCACCCAGCCCATGGCGAACCCCGGGCTGGGGCTGCTTCTGGCG
CTGGGCTGCGGTTCTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACAGACCACTTCTGC
AAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTGCTCCTGGCTGTGGGGCTG
GCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCAGTAGCGA
GGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCCATCTAGGTCCCCTCTCCTGCATCTGTCTCCCTTCATTGCTGTGTGACCTTGGGGAAA
GGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAGAAGGTACTTCAA
AGACTCTGCCCCTGAGGTCAAGAGAGGATGGGGCTATTCACTTTTATATATTTATATAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLPFLRLRWGRAWGQIQTTSANENSTVLPSSSTSSSDGNLRPEAITAIIVVFS
LLAALLLAVGLALLVRKLRKQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GCCAGGAATAACTAGAGAGGAACAATGGGGTTATTGAGAGGTTTTGTTTTCTCTTAGTTCTGTGCCTGCTGCACCAG
TCAAATACTTCCTTCATTAAGCTGAATAATAATGGCTTTGAAGATATTGTTCATTGTTATAGATCCTAGTGTGCCAGAA
GATGAAAAATAAATTGAACAAATAGAGGATATGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAGA
TTTTTTTTCAAAAATGTATCTATATTAATTCCTGAGAATTGGAAGGAAAATCCTCAGTACAAAAGGCCAAAACATGAA
AACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCCAGGTAGAGATGAACCATACACCAAGCAGTTCCACA
GAATGTGGAGAGAAAGCGAATACATTCACTTCACCCCTGACCTTCTACTTGGAAAAAACAAAATGAATATGGACCA
CCAGGCAAACTGTTTGTCCATGAGTGGGCTCACCTCCGGTGGGGAGTGTGATGAGTACAATGAAGATCAGCCTTTC
TACCGTGCTAAGTCAAAAAAATCGAAGCAACAGGTGTTCCGCAGGTATCTCTGGTAGAAAATAGAGTTTATAAGTGT
CAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAATTGATTCTACAACAAAACGTATGGAAAAGATTGTCAATTCTTT
CCTGATAAAGTACAAACAGAAAAGCATCCATAATGTTTATGCAAGTATTGATTCTGTTGTTGAATTTTGAACGAA
AAAACCCATAATCAAGAAGCTCCAAGCCTACAAAACATAAAGTGCAATTTTAGAAGTACATGGGAGGTGATTAGCAAT
TCTGAGGATTTTAAAAACACCATACCCATGGTGACACCACCTCCTCCACCTGTCTTCTCATTGCTGAAGATCAGTCAA
AGAATTGTGTGCTTAGTTCTTGATAAGTCTGGAGCATGGGGGGTAAGGACCGCCTAAATCGAATGAATCAAGCAGCA
AAACATTTCTGCTGCAGACTGTTGAAAATGGATCCTGGGTGGGGATGGTTCACCTTGATAGTACTGCCACTATTGTA
AATAAGCTAATCCAAATAAAAGCAGTGATGAAGAAACACACTCATGGCAGGATTACCTACATATCTCTGGGAGGA
ACTTCCATCTGCTCTGGAATTAATATGCATTTCAAGTGATTGGAGAGCTACATTTCCCAACTCGATGGATCCGAAGTA
CTGCTGCTGACTGATGGGAGGATAACACTGCAAGTTCTTGATTGATGAAGTGAACAAAAGTGGGGCCATTGTTTCAT
TTTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCAATTTTATGTT
TCAGATGAAGCTCAGAACAAATGGCCTCATTGATGCTTTTGGGGCTCTTACATCAGGAAATACTGATCTCTCCAGAAG
TCCCTTCAGCTCGAAAGTAAGGGATTAACTGAATAGTAATGCCTGGATGAACGACACTGTCTAATTTGATAGTACA
GTGGGAAAGGACACGTTCTTTCTCATCATGGAACAGTCTGCCTCCCACTATTTCTCTCTGGGATCCCACTGGAACA
ATAATGGAAAATTTACAGTGGATGCAACTTCCAAAATGGCCTATCTCAGTATTCAGGAACTGCAAGGTGGGCACT
TGGGCATACAATCTCAAGCCAAAGCGAACCAGAAACATTAACTATTACAGTAACCTTCTCGAGCAGCAAAATCTTCT
GTGCCTCCAATCACAGTGAATGCTAAAATGAATAAGGACGTAAACAGTTTCCCAGGCCAATGATTGTTTACGCAGAA
ATTCTACAAGGATATGTACCTGTTCTTGAGCCAATGTGACTGCTTTTCAATGAATCACAGAATGGACATACAGAAGTT
TTGGAACCTTTGGATAATGGTGCAGGCGCTGATTCTTTCAAGAATGATGGAGTCTACTCCAGGTATTTTACAGCATAT
ACAGAAAATGGCAGATATAGCTTAAAAGTTCCGGCTCATGGAGGAGCAACACTGCCAGGCTAAAATTACGGCCTCCA
CTGAATAGAGCCCGGTACATACCAGGCTGGGTAGTGAACGGGGAATTAAGCAAAACCCGCAAGACCTGAAATTGAT
GAGGATACTCAGACCACCTTGGAGGATTTGAGCCGAACAGCATCCGGAGGTGCATTTGTGGTATCACAAGTCCCAAGC
CTTCCCTTGGCTGACCAATACCCACCAAGTCAAAATCACAGACCTTGATGCCACAGTTTATGAGGATAAGATTATTCTT
ACATGGACAGCACCAGGAGATAATTTGATGTTGAAAAGTTCAACGTTATATCATAAGAATAAGTGAAGTATTCTT
GATCTAAGAGACAGTTTTGATGATGCTCTCAAGTAAATACTACTGATCTGTCAACCAAGGAGGCCAACTCCAAGGAA
AGCTTTGCAATTAACAGAAAATATCTCAGAAGAAAATGCAACCCACATATTATTGCCATTAAGTATAGATAAAA
AGCAATTTGACATCAAAAGTATCCAACATTGCACAAGTAACCTTTGTTTATCCCTCAAGCAAACTCTGATGACATTGAT
CCTACACCTACTCCTACTCCTACTCCTACTCCTGATAAAAGTCATAATCTGGAGTTAATATTCTACGCTGGTATTG
TCTGTGATTGGGTCTGTTGTAATTGTTAACTTTATTTAAGTACCACCATTTGAACCTTAACGAAGAAAAAATCTTC
AAGTAGACCTAGAAGAGAGTTTTAAAAACAAAACAATGTAAGTAAAGGATATTTCTGAATCTTAAATTCATCCCAT
GTGTGATCATAACTCATAAAAAATAATTTAAGATGTCGAAAAGGATACTTTGATTAAATAAAACACTCATGGATA
TGTAACAACTGTCAAGATTAAATTTAATAGTTTCATTTATTTGTTATTTTATTTGTAAGAAATAGTGATGAACAAAG
ATCCTTTTTCATACTGATACCTGGTTGTATATTATTGATGCAACAGTTTTCTGAAATGATATTTCAAATTGCATCAA
GAAATTTAAATCATCTATCTGAGTAGTCAAAATACAAGTAAAGGAGAGCAATAAACACATTTGGAAGAAAAA
AA

FIGURE 70

MGLFRGFVFLVLCLLHQSNSTFIKLNNGFEDIVIVIDPSVPEDEKIEQIEDMVTASTYLFE
ATEKRFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY
IHFTPDLLLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCISAGISGRN
RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVTEKASIMFMQSIDSVVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNSEDfKNTIPMVTPPPPVFSLLKISQRIVCLVLDKSGSMGGKDRLNR
MNQAACHFLLQTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK
YAFQVIGELHSQLDGSEVLLLLTDGEDNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG
SHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMNDTVIIDSTVGKDTFFFL
ITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR
AANSSVPPITVNAKMNDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA
GADSFKNMGVYSRYFTAYTENGRYSLKVRAGGANTARLKLRPPLNRAAYIPGWVVNGEIEANPP
RPEIDEDTQTTFLEDFSRITASGGAFVVSQVPSLPLPDQYPPSQITDLDTVHEDKIILTWTAPGDN
FDVGKVQRYIIRISASILDLRDSFDDALQVNTDLSPEANSKESFAFKPENISEENATHIFIAI
KSIDKSNLTSKVSINIAQVTLFIPQANPDDIDPTPTPTPTPTPKSHNSGVNISTLVLSVIGSVVI
VNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGGAACCCCTGGGAGTAGAGTACTGACAGCAAAGACCGGAAAGACCATACGTCCCGGGCAGGGGTGA
CAACAGGTGTCTATCTTTTGTATCTCGTGTGTGGCTGCCTTCTTATTCAAGGAAAGACGCCAAGGTAATTTTGACCCA
GAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACCCCAAGTTATGCCAGGATTACTAGAGAGTGTCA
ACTCAACAGCAAGCGGCTCCTTCGGCTTAACCTTGTGGTTGGAGGAGAGAACCCTTGTGGGGCTGCGTTCTCTTAGCA
GTGCTCAGAAGTGACTTGCCTGAGGGTGGACAGAAAGAAAGGTTCCCTCTTGTCTTGGCTGCACATCAGGAA
GGCTGTGATGGGAATGAAGGTGAAACTTGGAGATTTCACTTCAGTCATTGCTTCTGCCTGCAAGATCATCCTTTAA
AGTAGAGAAGTGTCTGTGTGGTGGTTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATGCAAGCAGCTC
CGGGGGCCCCAAACGCATGCTTCTGTGGTCTAGCCAGGGAAGCCCTTCCGTGGGGGCCCGGCTTTGAGGGATGCC
ACCGGTTCTGGACGCATGGCTGATTCCTGAATGATGATGGTTCCGCGGGGGCTGCTTGCCTGGATTTCCTGGGTGGTG
GTTTGTGCTGGTGTCTCTGTGTGTATCTGTCTGTCTGTACATGTTGGCCTGCACCCAAAGGTGACGAGGAGCAG
CTGGCACTGCCAGGCCCCAACAGCCACGGGGAAGGAGGGTACCAGGCCGTCTTCAAGAGTGGGAGGAGCAGCAC
CGCAACTACGTGAGCAGCTTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGAGCAGCTCAGG
AATGGGCAGTACCAAGCCAGCGATGCTGCTGGCTGGCTGGCTGGACAGGAGGCCCCAGAGAAAACCCAGGCCGACCTC
CTGGCCTTCTGTGCTCGCAGGTGGACAAGGCAGAGTGAATGCTGGCTCAAGCTGGCCACAGAGTATGACAGCAGTG
CCTTTCGATAGCTTTACTCTACAGAGGTGTACAGCTGGAGACTGGCCTTACCGGCCACCCGAGGAGAGGCTGTG
AGGAAGGACAAGCGGGATGAGTTGGTGGAGCCATTGAATCAGCCTTGGAGACCCTGAACAATCTTGCAGAGAACAGC
CCCAATCACCGTCTTACAGCGCCTCTGATTTCATAGAAGGGATCTACCGAACAGAAAGGGACAAAGGGACATTGTAT
GAGTCACTTCAAAGGGGACCAACAACAGGAATTAACAGGCTCATCTTATTTGACCATTCAGCCCCATCATGAAA
GTGAAAAATGAAAGCTCAACATGGCCACACGCTTATCAATGTTATCGTGCCTTAGCAAAAAGGGTGGACAAGTTC
CGGCAGTTCATGCAGAATTTCAAGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCTCACTGTTGTTTACTTTGGG
AAAGAAGAAATAAATGAAGTCAAAGGAATCTTGAAACACTTCCAAAGCTGCCAATTCAGGAACTTTACCTTCATC
CAGTGAATGGAGAATTTCTCGGGGAAAGGACTTGATGTTGGAGCCCGCTTCTGGAAGGGAAGCAACGCTCCTTCTC
TTTTCTGTGATGTGGACATCTACTTCACATCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAAG
GTATTTTATCCAGTCTTTTTCACTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTTGGAA
CAGCAGCTGGTCATAAAGAGGAACTGGATTTTGGAGAGACTTTGGATTTGGGATGACGTGTCACTGTTCTGGTTCAG
TTCAATCAATATAGTGGGTTTGTATCTGGACATCAAAAGCTGGGCGGAGAGGATGTGCACCTTTATCGCAAGTATCTC
CACAGCAACCTCATAGTGGTACGGACGCTGTGCGAGGACTCTTCCACCTCTGGCATGAGAAGCGCTGCATGGACGAG
CTGACCCCGGAGCAGTACAAGATGTGCATGCAGTCCAAGGCCATGAACGAGGATCCACGCGCCAGCTGGGCATGCTG
GTGTTACGACACGATAGAGGCTCACCTTCGCAACAGAAACAGAAAGCAAGTAGCAAAAAACATGAACCTCCAGA
GAAGGATTGTGGGAGACACTTTTCTTTTCTTTTCAATTAAGTGAAGTGGCTGCAACAGAGAAAAGACTTCCATAAA
GGACGACAAAAGAAATGGAGTGTGAGTGTGAGATGAGAAAGCCCTCCGATTCTCTCTGTTGGGCTTTTTACAACAGA
AATCAAAATCTCCGCTTTGCCGTGCAAAAGTAACCCAGTGTGACCCCTGTGAAGTGTCTGACAAAGGCAGAATGCTTGTG
AGATTATAAGCCTAATGGTGTGGAGGTTTGTATGGTGTTTACAATACACTGAGACCTGTTGTTTGTGTCTCATTTGA
AATATTCATGATTTAAGAGCAGTTTGTAAAAAATTCATTAGCATGAAAGGCAAGCATATTTCTCCTCATATGAATGA
GCCATCAGCAGGGCTCTAGTTTCTAGGAATGCTAAATATCAGAAGGCAGGAGAGGATAGGCTTATATGATACT
AGTGAGTACATTAAAGTAAATAAAATGGACAGAAAAGAAAAGAACCAATAATATCCTGTGTATATTTTCCCAAGAT
TAACCAAAAATAATCTGCTTATCTTTTTGGTGTCTTTTAACTGTCTCCGTTTTTTTCTTTTATTTAAAAATGCACT
TTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTACCACTTTGCAAGCCTTACAAGAGAGCACAAGTTGGCCTAC
ATTTTATATTTTTAAGAAGATACTTTGAGATGCATTATGAGAACTTTCACTTCAAAGCATCAAATGATGCCATAT
CCAAGGACATGCCAAATGCTGATTTCTGTGAGGCACTGAATGTGAGGCTTGAACATAGGGAAGGAATGGTTGTACT
AATACAGACGTACAGATACTTTCTCTGAAGATATTTTCAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATGAC
ACTTTCTGCTTTACAGAAAAGGAACTCATTGAGACTGGTGATATCGTGATGTACCTAAAAGTCAGAAAACCATTTTT
CTCCTCAGAAGTAGGGACCGCTTTCTACCTGTTTAAATAAACCAAGTATACCGTGTGAACCAACAATCTCTTTTC
AAAACAGGGTGTCTCTCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAATATATATATATATATATATTTGT
GAAAGATCAATCCATCTGCCAGAATCTAGTGGGATGGAGTTTGTGTACATGTTATCCACCCAGGCCAGGTGGAAG
TAACTGAATTATTTTTTAAATTAAGCAGTTCTACTCAATCACCAGATGCTTCTGAAAATTGCATTTTATTACCATTT
CAAACATTTTTTAAAAATAAATACAGTTAACATAGAGTGGTTTCTTCATTATGTGAAAATTATTAGCCAGCACCAG
ATGCATGAGCTAATTTCTCTTTGAGTCTTGTCTGTTTGTCTCACAGTAAACTCATTGTTTAAAGCTTCAAGAAC
ATTCAAGCTGTTGGTGTGTTAAAAATGCATTGTATTGATTGTACTGGTAGTTTATGAAATTAATTAACACAGG
CCATGAATGGAAGGTGATTTGCACAGCTAATAAAATGATTTGTGGATATGAA

FIGURE 72

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQWE
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALETLNPA
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVKNEKLNMAN
TLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPVLSQY
NPGIYGHHDVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDLDIKWGAGEDVHLYR
KYLHSNLIVVRTPVRGLEHLWHEKRCMDLTPEQYKCMQSKAMNEASHGQLGMLVFRHEIEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

[illegible][illegible]

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHPASYPECGNNAQSPIDIQTDSVTFDPDLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQINSEAT
FAELHIVHYDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCCCTTGGGGACGGGCAGTTCCTGTGTCT
TCTGGTGGTTTGCTTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGTCCTACA
ATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTTCATCACAA
ATTGGCCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGTCTGACAGCTCC
AGAGAAGTGGAAAGAGAAATCCAGAAGACCTTCTGTTTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCAC
ACGCTGGTGTCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCTTCGTCCC
AGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCAT
CAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGTGTTCCTTTT
TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAGCAAATTT
GATTTTGATTATGGAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGAAAAATCGTGATTA
ACTTTATCACCTCAATATCTCGGATGATTCATAAATTTCTCATCAGGATATGAGTTTACTGGGA
AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGA
GGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGAAATTTTTGTGACTCTGAAG
AAAACACGGAAGGTACTTCTCTCACCCAGCAAGAGTCCCTCAGCAGAACAAATACCCCGGATAAA
ACAGTCATTGAATATGAATATGATGTCAGAACCACTGACATTTGTGCGGGGCTGAAGAGCAGGA
GCTCAGTTTGCAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAG
TCTTGGGCCCCGAAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGACCCCTGGCG
CAGGAGCACACAGACTCGGAGGAGGGGCCGGAGGAAGAGCCATCGACGACCTGGTCTGACTGGGA
TCCCCAAACTGGCAGGCTGTGTATTCCTTCGCTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCG
AGCCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCT
CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGTTATA
TGTGCAGATGGAAAACTGATGCCAAACACTTCCTTTTGCCTTTTGTTCCTGTGCAAACAAGTGAG
TCACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTCCAGTTTGTGAGTGT
CTGTGAGAATTACTTATTTCTTTTCTCTATTCTCATAGCACGTGTGTGATTGGTTCATGCATGTA
GGTCTCTTAACAATGATGGTGGGCTCTGGAGTCCAGGGGCTGGCCGGTTGTTCTATGCAGAGAA
AGCAGTCAATAAATGTTTGCCAGACTGGGTGCAGAATTTATTCAGGTGGGTGT

FIGURE 76

MSYNGLHQRVFKELKLLTLCSISSQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY
SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRRAPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNFEFDRFFVPAEK
IVINFITLNISSDDSKISHQDMSLLGKSSDVSSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFC
DSEENTEGTSLTQQESLSRTIPDPKTVIEYEYDVRTTDCAGPEEQELSLQEEVSTQGTLLESQA
ALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQDS
EGCEPSEGDGLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGACAC
CTGGGAAGATGGCCGGCCCCGTGGACCTTCACCCCTTCTCTGTGGTTTGCTGGCAGCCACCTTGATC
CAAGCCACCCTCAGTCCCACTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCCTGAAGCACATC
ATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTCGGCCAATGACCA
GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
TCGTGGAGTTCACATGACGACTGAGGCCCAAGCCACCATCCGCATGGACACCAGTGCAAGTGGC
CCCCCGCCTGGTCTCAGTACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAAGTCTGTGA
TAAGCTCTCCTTCTGGTGAACGCCTTAGCTAAGCAGGTGATGAACCTCCTAGTGCCATCCCTGC
CCAATCTAGTGAAAAACCAGCTGTGTCCCGTGATCGAGGCTTCCTTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCCATTTCCCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTA
TCCTGCCATCAAGGGTGACACCATTAGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAA
AGGTGACCAAGTGGTTCAATAACTCTGCAGCTTCCTTGACAATGCCACCCTGGACAACATCCCC
TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGA
ATTCATGGTCCTGTTGGACTCTGTGCTTCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGC
TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGAC
ACTCCCGAGTTTTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTT
TCCCTCCAGTGAAGCCCTCCGCCCTTTGTTACCCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGT
TTTACACCAAAGGTGACCAACTTATACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTG
ATGAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTGAAGG
CCTTGGGATTCGAGGCAGCTGAGTCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCCTCC
TTGTGGAAACCCAGCTCTCCTGTCTCCAGTGAAGACTTGGATGGCAGCCATCAGGGAAGGCTGG
GTCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACACTTG
CCTGTGAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVKEKLTQELKDNATSILQQLPLLSAMREK
PAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDITQLYLGAALLDSQGVKVT
KWFNNSAASLTMTPLDNIPFSLIVSQDVVKAAVAVALSPPEEFMVLLDSVLPESAHLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEALRPLFTLGEASSEAQFYT
KGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIHSILLPNQNGKLRSQVPSLVKALG
FEAAESSLTGDALVLTTPASLWKPSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACCTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGTTGCCAT
GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAGCAGTTGGCTTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCCCTTCTGGGCCTGCCCGCTGACATCCAGGCTGCCAGGCCATGATGGTGACATCCAGTGAAT
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTTCATCCTTGAGGGCCTCCTGGGATTC
ATTCTGTTCCTGGAATCTTCATGGGATCCTACGGGACTTCTACTCACCAGTGGTGCCTGACAG
CATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTCTTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTTTCTGCTCATCCCAGAGAAATCGCTCCAACCTACTACGATGCCTAC
CAAGCCCCAACCCTCTTGCCACAAGGAGCTCTCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGA
GTTCAATTCCTACAGCCTGACAGGGTATGTGTGAAAGAACCAGGGGCCAGAGCTGGGGGGTGGCTG
GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGG
GCTAGTGTAAACAGCATGCAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCC
TCACCTTGCTGCTCCCCCTGCCCTAAGTCCCCAACCCCTCAACTTGAAACCCCATTCCTTAAGCCA
GGACTCAGAGGATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACA
TCCCACTGACTGACCCTCTGTGATCAAAGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATT
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTC
CCTCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCACTCTTGTTATGACTCCACAGTGTCCA
GACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

FIGURE 8o

MASLGLQLVGYYLGLLGLLGLTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMVTSSAIISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAWNLHGILRDFYSPVLPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKSEFNYSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCTTTC
CCCGCGTTCTCTTTCCACCTTTCTCTTCTTCCCACCTTAGACCTCCCTTCCCTGCCCTCCTTTCCT
GCCCACGCTGCTTCTGGCCCTTCTCCGACCCGCTCTAGCAGCAGACCTCCTGGGGTCTGTGG
GTTGATCTGTGGCCCTGTGCCCTCCGTGTCTTTTCGTCTCCCTTCCCTCCCGACTCCGCTCCCGG
ACCAGCGGCTGACCCTGGGGAAAGGATGTTTCCCAGGTGAGGGTCTCTCCTCCTTGTCTGGGA
CTCGCGCTGCTCTGGTTCCCCCTGGACTCCACGCTCGAGCCCGCCAGACATGTTCTGCCTTTT
CCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTTGAGCCACAAGGCCTGATGT
ACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTGTTACCGCCTCCACTGTCCGCCT
GTCCACTGCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCACAC
TCCCTCTGGACTCCGGGCCCCACCAAGTCTGCCAGCACAACGGGACCATGTACCAACACGGAG
AGATCTTCAGTGCCCATGAGCTGTTCCCTCCCGCCTGCCCAACCAGTGTGTCTCTGCAGCTGC
ACAGAGGGCCAGATCTACTGCGGCCTCACAACCTGCCCGAACCAGGCTGCCAGCACCCCTCCC
ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAAGATCGGATGAAGAGGACA
GTGTGCAGTCGCTCCATGGGGTGAGACATCTCAGGATCCATGTTCCAGTGATGCTGGGAGAAAG
AGAGGCCCCGGGCACCCAGCCCCACTGGCCTCAGCGCCCTCTGAGCTTCATCCCTCGCCACTT
CAGACCCAAGGGAGCAGGCAGCACAACCTGTCAAGATCGTCCTGAAGGAGAAACATAAGAAAGCCT
GTGTGCATGGCGGGAAGACGTACTCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGC
CCCTTGCCCTGCATCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGTGCAAGATTTGCCAGAGG
ACAAAGCAGACCCTGGCCACAGTGAATCAGTTCCTACAGGTGTCCCAAGGCACCGGGCCGGGTC
CTCGTCCACACATCGGTATCCCCAAGCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGGC
CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAACTGAGGCTCAGAGAG
GTGAAGTACCTGGCCAAGGCCACACAGCCAGAATCTTCACTTGACTCAGATCAAGAAAGTCAG
GAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCACGAAGGTCCT
GGAACGTCTTCTAGCCCAGACCTGGAGCTGAAGGTCACGGCCAGTCCAGACAAAGTGACCAAG
ACATAACAAAGACCTAACAGTTGCAGATATGAGCTGTATAATTGTTGTTATTATATATTAATAAA
TAAGAAGTTGCATTACCCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPD MFCLFHGKRYSPGESWHPYLEPQGLMYCLRTCSE
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKC VEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELF
PSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPPTAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKHKKACVHGGKTYS
HGEVWHPAFAFAFGPLPCILCTCEDGRQDCQRVTCPT EYPCRHPKVAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLDSDQESQEARLPERG TALPTARWP PRRSLERLPSPDPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGGCCCTCCGCTCACGCAGAGCCTCTCC
 GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTTCTCTCTAATCCATCCGTCACCTCTCCTGTCA
 TCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTTGGTTCTGAGTC
 TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTTGGGCCAGACAAGCCTGTCCAGGCCTTGGTGGGGGAG
 GACGCAGCATTCTCCTGTTTCTGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTCTTTCAGGGG
 CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATC
 AAGGCAGGACAAAACCTGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTCTGAGGCTGGAAAACATTACT
 GTGTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCAGTCTTACTACCAGAAGGCCATCTGGGAGCT
 ACAGGTGTGAGCACTGGGCTCAGTTCTCTCATTTCATCACGGGATATGTTGATAGAGACATCCAGCTAC
 TCTGTCTCAGTCTCGGGCTGGTTCCCCCGGCCACAGCGAAGTGGAAGGTCCACAAGGACAGGATTTGTCC
 ACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTTGATGTGGAGATCTCTCTGACCGTCCAAGAGAA
 CGCCGGGAGCATATCCTGTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
 GAGATACCTTTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGGCCTA
 TTTTTTGGCATTGTTGGACTGAAGATTTTCTTCTCCAAATTCAGTGGAATCCAGGCGGAACCTGGACTG
 GAGAAGAAAGCACGGACAGGCAGAAATTGAGAGACGCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAG
 AGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAACCTGTAACCCATAGAAAAGCTCCCCAGGAGGTG
 CCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTCTCAGAGTTTCCAAGCAGGGAAACATTA
 CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGCTGGGAGTGTGCCGGGATGATGTGGACAGGAGGA
 AGGAGTACGTGACTTTGTCTCCCGATCATGGGTACTGGGTCTCAGACTGAATGGAGAACATTTGTATTTT
 ACATTAAATCCCCGTTTTATCAGCGTCTTCCCCAGGACCCACCTACAAAATAGGGGTCTTCTGGACTA
 TGAGTGTGGGACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCCTGACATGTCGGTTTG
 AAGGCTTATTGAGGCCCTACATTGAGTATCCGTCCTATAATGAGCAAAATGGAATCCCATAGTCATCTGC
 CCAGTCACCCAGGAATCAGAGAAAGAGGCCCTCTTGGCAAAGGGCCTCTGCAATCCAGAGACAAGCAACAG
 TGAGTCTCTCTCACAGGCAACCACGCCCTTCTCCCCAGGGGTGAAATGTAGGATGAATCACATCCCACAT
 TCTTCTTTAGGGATATTAGGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGCCGGCCAAGGTGGCTTCCA
 GATGAAGGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTCTGCTGCCCTGAGCTGGGAGGGAAGAAGG
 CTGACATTACATTTAGTTTGTCTCTCACTCCATCTGGCTAAGTGATCTTGAATACCACCTCTCAGGTGAAG
 AACCCTCAGGAATCCCATCTCACAGGCTGTGGTGTAGATTAAGTAGACAAGGAATGTGAATAATGCTTAG
 ATCTTATTGATGACAGAGTGTATCCTAATGGTTTGTTCATTATATTACACTTTCAGTAAAAAAA

FIGURE 84

MALMLSLVLSLLKLGSGQWQVFGPDKPVQALVGEDAAAFSCFLSPKTNAEAMEVRFFRGQFSSVVH
LYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISQSYQKAIWELQ
VSALGSVPLISITGYVDRDIQLLCQSSGWFPRTAKWKGPQGDLSTDSRTNRDMHGLFDVEISL
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAELDWRRKHGQAEIRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF
TRKSVVASQSFGAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYWVLRNLNGEHLFT
LNPRFISVFPRTPTTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
PIVICPVTQESEKEASWQRASAIPESTSSESSSQATTPLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCAGACATGCTGCTGCTGCTGCTGCCCCCT
GCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAACTGCTGACGATGCAGAGTTCCGTGA
CGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGCTGGATT
TACCCTGGCCAGTAGTTTCATGGCTACTGGTTCCGGGAAGGGGCCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTTG
GGGACCCACATACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGAAGAAGTGATGCGGGGAGA
TACTTCTTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCAACATCCTCATCCCAGGCACCCTGGAGTCCGGCTGCCCCC
AGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGGACACCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCGGTGCTCACCTCATCCCACA
GCCCCAGGACCATGGCACCAGCCTCACCTGTGAGGTGACCTTCCCTGGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTCAACGTGTCTACCCGCCCTCAGAACTTGACCATGACTGTCTTCCAAGGA
GACGGCACAGTATCCACAGTCTTGGGAAATGGCTCATCTCTGTCACTCCAGAGGGCCAGTCTCT
GCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCCTGACCTGTGCCCTCACAGCCCTCAAACCCGGGGGTGCTGGAGCTGCCTTGGGTGCAC
CTGAGGGATGCAGCTGAATTACCTGCAGAGCTCAGAACCTCTCGGCTCTCAGCAGGTCTACCT
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTCCGGGGAGCTGGAG
CCACAGCCCTGGTCTTCTGTCTTCTGCGTCATCTTCGTTGTAGTGAGGTCTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTGAGGGGTTT
AGCCTCTCAGGGGCCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCCAG
CTTCTGCCGCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG
AAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATCAGAAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCCTCCAGGCAAGGGAGAAGTCA
GAGGCTGATCTTGTAGAATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCCC
TCCCTTTTATTTTTTAACTAAAAGACAGACAAATTCCTA

FIGURE 86

MLLLLLP LLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSYPSHGWIYPGPVVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDREHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCVWPWACEQGTPPMISWIGTSVSPLDPSTTRS
SVLTLIPQPQDHGTS LTCQVTFPGASVTTNKTVHLNVSYPPQNLMTVFQGDGTVSTVLGNGSSL
SLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLVNVLQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGDGTGIE
DANAVRGSASQGPLEPWAEDSPDPQPPASARSSVGEGLQYASLSFQMKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTAC
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTGTTTC
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGTGCATTTGA
TGGCCTGTATTTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGACATGACCTCTG
GGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAACTGGGC
CAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCCAATAAGTCCCCCATGCAGCACTGG
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCT
GTTTGGCATCTACCAGAAATATCCAGTGAAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC
CGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCAGAAAACAGCATCTTATTACTCACCTAT
GGCCAGCGGGAATTCAGTGCGGGATTTGTTAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAA
CGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
GATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATAT
GGAACATCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTTGTGGGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAA
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTILVASVHENDMRGKCTVGDWSSQQGSKADYPEGDGNWANYNTFGSAEAAT
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTDNGPVIPIVVYDFGDAQKTASYISPYGQREFTAGFVQFRVENNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSRSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGAGAC
CGCCGCCCTTGTCCCGAGGGCCATGGGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCTGACG
CTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACATACAGGCCTG
CCTGCCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTCAGCTGGTGGCCGCGCTCT
CTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCTCTCAGGAGTCTCCATGTTCAAC
AGCACCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCTGTCCTTCTTCAT
ATTCGAGCGTTGGGAGTGCACGTATTGGTACATTTTGTCTTCTGCAGTGCCCTTCCAGCTG
TCACTGAAATGGCTTTATTTCGTCACCGTCTTTGGGCTGAAAAAGAAACCCTTCTGATTACCTTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGTATTCTGGAAGAAGGAAG
GCATAGGCTTCGGTTTTCCCTCGGAACTGCTTCTGCTGGAGGATATGTGTTGGAATAATTACG
TCTTGAGTCTGGGATTATCCGCATTGTATTTAGTGCTTTGTAATAAAATATGTTTTGTAGTAACA
TTAAGACTTATATACAGTTTTAGGGGACAATTAAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCALPAVTEMALFV
TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGAC
TCGCTGCTGCTTCGTGTTCCCTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG
GTCTCCCGGCTCAGAGGACCCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCT
CGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATTCCACTCTCCTAGGGCTGCT
GGCCCCGCTTGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCCGAACCACAGCCCCC
CACCCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAACATCAAGACGGTG
GCCCTGAACCTGCTCGTCACAGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAACATCTCCATCAGCCTCGTGCCCCCAGTAAAGCTGTAG
AGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAAATCTTCAACTGCCGGATG
GAGTGGGAGAAGGTAGAACGGGGCCCGCGGACCTCGCTTTCACCCACGACCCAGCCAAGATCTG
CTCCCGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCAGCCCTTCAAAGTCGTCTGTG
TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAACCTAC
CATAGTGATACCCCCCTACTACCCATCTGGGTGACCCGGGGCAGGCCACAGAGGCCAGGCCAGGGC
TGGAAGGACAGGCCTGCCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAGGGGTGGGCCCTC
AGGCAGGGAGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG
AGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG
AGGAGGAGTGGGCTCTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGG
TCCCCGAGGCCTGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGC
CCTTGTTCTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTTCATCAGTGTGGACA
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGGTGGG
GCCAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCTTGTCTGTGCTGAGCATGG
CATGAGGCTGAAGTGGCAACCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
CAGGCCACCCCTTTCCAAAATTCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATG
GCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCTCCACACTAAGGCCACAGCCCATC
CGCGTGCTGTGTGTCCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCG
GAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTCAGACCGGGGTCTCCCGGATCTGGATGGCGC
CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCGGCCGAGAGCATGTGCTGGATCTGTTC
TGTGTGTCTGTCTGTGGGTGGGGGAGGGGAGGGAAGTCTGTGAAACCGCTGATTGCTGACTTT
TGTGTGAAGAATCGTGTCTTGGAGCAGGAAATAAGCTTGCCCCGGGGCA

FIGURE 92

MQLTRCCFVFLVQGSILYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPFKVVVCVYIAFYSTDYRLVQKVC PDYNYHSDTPYYP SG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTCGCC
CTTTATGTCTTCACCATCGCCATCGAGCCGTGCGTATCATCTTCCTCATCGCCGAGCTTTCTT
CTGGTTGGTGTCTCTACTGATTTGCTCCCTGTTTGGTTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTGTCTATATCCAAGAA
ATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTTGAAGAGTATAAACCC
AGGTGAGACAGCACCCCTCTATGCGACTGCTGGCCTATGTTTCTGGCTTGGGCTTTGGAATCATGA
GTGGAGTATTTTCCTTTGTGAATACCCCTATCTGACTCCTTGGGGCCAGGCACAGTGGGCATTCAT
GGAGATTCTCCTCAATTCTTCCTTTATTCAGCTTTTCATGACGCTGGTCATTATCTTGCTGCATGT
ATTCTGGGGCATTGTATTTTTTGTGAGTGTGAGAAGAAAAGTGGGGCATCCTCCTTATCGTTC
TCCTGACCCACCTGCTGGTGTGAGCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCG
TCAGCATTTATAATCCTGGTGCTCATGGGCACCTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCG
AAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGAAGTTTCTTCTTTACAACCAGCGCTCCAGAT
AACCTCAGGGAACCAGCACTTCCCAAACCGCAGACTACATCTTTAGAGGAAGCACAACTGTGCCT
TTTTCTGAAAATCCCTTTTTCTGGTGGAATTGAGAAAGAAATAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWEMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLPGTGVGIHGDSPOFFLYSAFMTLVIILLHVFWGIVFFDGCEKKKWGILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSCLKCLLCQDKNFLLYNQRSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTTCACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTTGTCTTCTAGGATCAAC
TCGGTCATTACCACAGCTCAAACCTGCTTTGGGACTCCCTCCCACAAACTGGCTCCGGATCAGG
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTTCCTTCTTTAAGTCTGATACCATTAACA
CAGATGCTCACACTGGGGCCAGATCTGCATCTGTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTTGTGCACACAACCTTGAGCCCAGGGCAGTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCATCCATTCCCTTGTTCCCGGGAGGCATCCTGCCCACCAGTCAGGCAGGGGC
TAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAGGAGCAGGTGTAAATCCTGCCACCCAGG
GAACCCAGCAGGCCCGCCTCCCAACTCCCAGTGGCACAGATGACGACTTTGCAGTGACCACCCCT
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATCA
GTAAGCTGTTTTCAAATTTTTTCAACTAAGCTGCCTCGAATTTGGTGATACATGTGAATCTTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAATTCTTAATT
TACCTGAAAATATTCTTGAAATTTAGAAAATATGTTCTATGTAGAGAATCCCAACTTTTAAAAA
CAATAATTCAATGGATAAATCTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAACATATTTGAAAAC TGAAA
AAA

FIGURE 96

MRSTIILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQSQSNQVFPSSLIPLTQM
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCTTGCCTTGCCCCACCCAGCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT
 CTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCCTGACCATG
 GTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCAGGCTCTCCCCAAGGCCAGCCTGC
 AGAGCTGTCTGTGGAAGTTCAGAAAACATATGGTGGAAATTTCCCTTTATACCTGACCAAGTTGC
 CGCTGCCCCGTGAGGGGGCTGAAGGCCAGATCGTGTGTGAGGGGACTCAGGCAAGGCAACTGAG
 GGCCCATTTGCTATGGATCCAGATTCTGGCTTCCGTGTGTGAGGAGGGCCCTGGACCCAGAGGA
 GCAGGCAGAGTACCAGCTACAGGTACCCCTGGAGATGCAGGATGGACATGTCTGTGGGGTCCAC
 AGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCCATTTCTCTCAAGCCATCTAC
 AGAGCTCGGCTGAGCCGGGTACCAGGCTGGCATCCCCCTTCTCTTCTTGGGCTTCAGACCG
 GGATGAGCCAGGCACAGCCAACCTCGGATCTTCGATTCCACATCTGAGCCAGGCTCCAGCCCAGC
 CTTCCCCAGACATGTTCCAGCTGGAGCCTCGGCTGGGGGCTCTGGCCCTCAGCCCCAAGGGGAGC
 ACCAGCCTTGACCACGCCCTGGAGAGGACCTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGA
 CCAGGCCTCAGGCCACCAGGCCACTGCCACCGTGGAAAGTCTCCATCATAGAGAGCACCTGGGTGT
 CCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCTTATACCCGCACACATGGCCCCAGGTA
 CACTGGAGTGGGGTGTATGTGCACTATCACCTGGAGAGCCATCCCCGGGACCCCTTGAAGTGAA
 TGCAGAGGGAAACCTCTACGTGACCAGAGAGCTGGACAGAGAAGCCCAGGCTGAGTACCTGCTCC
 AGGTGCGGGCTCAGAATTCCTATGGCGAGGACTATGCGGCCCTCTGGAGCTGCACGTGCTGGTG
 ATGGATGAGAATGACAACGTGCCTATCTGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCT
 CAGTCCACCAGGTACTGAAGTGACTAGACTGTGAGCAGAGGATGCAGATGCCCCCGGCTCCCCCA
 ATTCCACGTTGTGTATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTC
 CAGGTGGACCCCACTTCAGGCAGTGTGACGCTGGGGGTGCTCCCACTCCGAGCAGGCCAGAACAT
 CCTGCTTCTGGTGCTGGCCATGGACCTGGCAGGCGCAGAGGGTGGCTTCAGCAGCACGTGTGAAG
 TCGAAGTCGCAGTCACAGATATCAATGATCAGCCCCCTGAGTTCATCACTTCCCAGATTGGGCCCT
 ATAAGCCTCCCTGAGGATGTGGAGCCCCGGGACTCTGGTGGCCATGCTAACAGCCATTGATGCTGA
 CCTCGAGCCCCGCTTCCGCCCTCATGGATTTTGCATTGAGAGGGGAGACACAGAAGGGACTTTTG
 GCCTGGATTGGGAGCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG
 GCAGCTCCAAGTCATGAGGTGGTGGTGGTGGTGCAGAGTGTGGCGAAGCTGGTGGGGCCAGGCC
 AGGCCTGGAGCCACCGCCACGCTGACTGTGCTAGTGGAGAGAGTGTGCCACCCCCAAGTTGG
 ACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTGGCCAGCCGGCTCTTCTGCTGACCATC
 CAGCCCTCCGACCCCATCAGCCGAACCCCTCAGGTTCTCCCTAGTCAATGACTCAGAGGGCTGGCT
 CTGCATTGAGAAATTTCTCCGGGGAGGTGCACACCGCCAGTCCCTGCAGGGCGCCAGCCTGGGG
 ACACCTACACGGTGCTTGTGGAGGGCCAGGATACAGCCCTGACTCTTGCCCTGTGCCCTCCCAA
 TACCTTGCACACCCCGCCAAGACCATGGCTTGATCGTGAGTGGACCCAGCAAGGACCCCGATCT
 GGCCAGTGGGCACGCTCCCTACAGCTTACCCCTTGGTCCCAACCCACCGGTGCAACGGGATTGGC
 GCCTCCAGACTCTCAATGGTTCCCATGCCTACCTCACCTTGGCCCTGCATTGGGTGGAGCCAGT
 GAACACATAATCCCCGTGGTGGTCAGCCACAATGCCAGATGTGGCAGCTCCTGGTTGAGTGAT
 CGTGTGTGCTGCAACGTGGAGGGGAGTGCATGCGCAAGGTGGGCGCATGAAGGGCATGCCCA
 CGAAGCTGTGCGCAGTGGGCATCCTTGTAGGCACCCCTGGTAGCAATAGGAATCTTCTCATCCTC
 ATTTTCACCCACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTGCCCT
 GAAGGCGACTGTCTGAATGGCCAGGCAGCTCTAGCTGGGAGCTTGGCCTCTGGCTCCATCTGAG
 TCCCTGGGAGAGAGCCAGCACCCAAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCTCCA
 TCTGCCCTGGGGTGGAGGCACCATCACCATCACCAGGCATGTCTGCAGAGCCTGGACACCACTT
 TATGGACTGCCCATGGGAGTGTCTCAAATGTGAGGTGTTTGCCCAATAATAAAGCCCCAGAGAA
 CTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSVPAQALPKAQPaelSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQPSDFQLEPRLGALALSPKG
STSLDHALERTYQLLVQVKMDGQASGHQATATVEVSIIESTWVSLEPIHLAENLKVLYPHHMAQ
VHWSGGDVHYHLESHPPGPFVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDIAAPLELHVL
VMDENDNVPICPPRDPVSIPELSPPGTEVTRLAEDADAPGSPNSHVYQLLSPEPEDGVEGRA
FQVDPTSGSVTLGVLPLRAGONILLVLAMDLAEGGFSSTCEVEVAVTDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTADLEPAFRIMDFAIERGDTEGTFGLDWEPSGHVRLRLCKNLSY
EAAPSHEVVVVQSVAKLVGPGPGGATATVTVLVERVMPPPKLDQESYEASVPISAPAGSFLLT
IQPSDPISTRFSLVNDSEGLWCIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALTLPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGNPTVQRDWRLQTLNGSHAYLTALHWVEP
REHIIPVVVSHNAQMWQLLVRVIVCRNVEGQCMRKVGRMKGMPTKLSAVGILVGTTLVAIGIFLI
LIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCCACGCCTG
AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTTA
TCTCTTCACCTTCAAGTCCCCTTTCTCAAGAATCCTCTGTTCTTTGCCCTCTAAAGTCTTGGTAC
ATCTAGGACCCAGGCATCTTGCTTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
TTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAAATTCCAATGAGACTAGCACC
TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG
TGTGACCTCCAGTGGGGTCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGGTCA
GCATAGTCACCAACTCTGAGTTCCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCTCCAGTGGGGCCAGCACAGTCAACCACT
CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC
AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGGCCAGCACAGCCAC
CAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT
GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCCAATGGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCCA
GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAG
TCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG
GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCGGCACAGCCACCAACT
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCAACCAATTCTGAGTCCAGCACACCCTCC
AGTGGGGCCAAACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCAAACAGCCAC
CAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGGTCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACA
GCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTCAACCAATTCTGAGTCCAGCACAACTCCAGTGGGGCCA
ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA
ATGCACACAACTTCCCATAGTGCATCTACTGCAGTGTGAGGCAAGCCTGGTGGGTCCCTGGT
GCCGTGGGAAATCTTCCATCACCTGGTCTCGGTGTGGCGGCCGTGGGGCTCTTTGCTGGGC
TCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGTCTACCACCT
CATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCCACAGGCCAG
GTGGAGTCCCTAACTGGTTCTGGAGGAGACCAGTATCATCGATAGCCATGGAGATGAGCGGGAGGA
ACAGCGGGCCCTGAGCAGCCCCGGAAGCAAGTGCCGCATTCTTCAGGAAGGAAGAGACCTGGGCA
CCCAAGACCTGGTTTCCTTTTCAATTCATCCAGGAGACCCCTCCAGCTTTGTTTGAGATCCTGAA
AATCTTGAAGAAGGTATTCTCACCTTTCTTGCCTTTACCAGACACTGGAAAGAGAATACTATAT
TGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACACGACAAAGAGAAGCTGTGCTTG
CCCCGGGTGGGTATCTAGCTCTGAGATGAATCAGTTATAGGAGAAAACCTCCATGCTGGACTC
CATCTGGCATTCAAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKMQKGNVLLMFGLLLHLEAATNSNETSTSTANTGSSVSSGASTATNSGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNESSTSSGASTATNESSTPSS
GASTVTNSGSSVTSSGASTATNESSTVSSRASTATNESSTLSSGASTATNSDSSTSSGASTA
TNSESSTSSGASTATNESSTVSSRASTATNESSTSSGASTATNESRRTSNGAGTATNES
STTSSGASTATNSDSSTVSSGASTATNESSTSSGASTATNESSTSSGASTATNSDSSTSS
GAGTATNESSTVSSGISTVTNSESSTPSSGANTATNESSTSSGANTATNESSTVSSGASTA
TNSESSTSSGVSTATNESSTSSGASTATNSDSSTTSSEASTATNESSTVSSGISTVTNSES
STTSSGANTATNSGSSVTSAGSGTAALTMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVVA
AVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGPGPGGNHGAPHRPRWSPNWFRRPVSSI
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTTCGCGACGGAGGTTGTGACCCCTA
CGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGCGGCGTGCCCTGCTTGTCACAGGTG
GGAGGCTGGAACATCAGGCTGAAAAACAGAGTGGGTACTCTCTTCTGGGAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGGAAGGAAATTGTGGTGCTTCTGAACCCATGGTCAATT
AACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTTGGAATCAT
GGTGTGTCATGGAAAGGATTACTTTTATACTGACTCTGTTTTGGGGAAGCTTTTTTGAAGCATTT
TCATGCTGAGTCCCTTTTACCTTTGATGTTTGTAACCCATCTTGGTATCGCTGGATCAACAAC
CGCCTTGTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTTGGTGTAAGT
GATTATAACTGGGGATGCATTTGTTCTCGGAGAAAGAAGTGTCTATTATCATGAACCATCGGACAA
GAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCCTCAAAGCGAGTCTCAAAGGTGTTCTCGGATTTGGTTGGGCCATGCAGGCTGCTGCCTATAT
CTTCATTATAGGAAATGGAAGGATGACAAGAGCCATTTGAAGACATGATTGATTACTTTTGTG
ATATTACGAACCACTTCAACTCCTCATATCCGAGAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTCAGAAATATGAATATGTTTTACATCCAAG
AACTACAGGCTTTACTTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTCACACATTCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTTT
CCCAGGGAAATCCACTTTACGTCCACCGGTATCCAATAGACACCCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGTGCCACAAACGGTGGGAAGAGAAAGAGAGGCTGCGTTCCTTCTATCAAG
GGGAGAAGAATTTTTATTTTACCGGACAGAGTGTCTTCCACCTTGCAAGTCTGAACTCAGGGTC
CTTGTGGTCAAATTGCTCTCTATACTGTATTGGACCCTGTTTACGCCCTGCAATGTGCCTACTCAT
ATATTTGTACAGTCTTGTTAAGTGGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGA
GAATATTTGGTGGACTGGAGATCATAGAACTTGCATGTTACCGACTTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAAATGAGTAAGATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTTG
GAAATGTTCTAAACCTTTCTAAGCTCAGATGCATTTTTGCATGACTATGTGCAATATTTCTTACT
GCCATCATTATTTGTTAAAGATATTTTGCCTTAATTTTGTGGGAAAAATATTGCTACAATTTTT
TTAATCTCTGAATGTAATTCGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTT
GGGCCAGAATATTATTAAACAATCATCAGGCTTTTAAA

FIGURE 102

MHSRGREIVLLNPWSINEAVSSYCTYFIKQDSKSGIMVSWKGIYFILTLFWGSFFGSI FMLSP
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
FLWNCIMRYSYLRLKICLKASLKGVPFGGWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSKSRSNFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLQGDFFPREIHFHVHRYPIDTLPTSKEDLQLWCHKRWEEKEERLRSFYQGEKNF
YFTGQSVIPPCKSELRLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
LEIIEIACRYRLHKKPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGACC
TCCAAATCATCCATCCACCCCTGCTGTCATCTGTTTTTCATAGTGTGAGATCAACCCACAGGAATA
TCCATGGCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACAGTGGCAAGT
CACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACGCCGTGTTCTCTGCTCCCTCT
TTCCTGAGACCAAGTGCAGAGGCTATGGAAGTGGGTTCTTCAGGAATCAGTTCCATGCTGTGGTC
CACCTCTACAGAGATGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCGAGGGAGAAGTGA
GTTTGTGAAGGACTCCATTGCAGGGGGGCGTGTCTCTAAGGCTAAAAACATCACTCCCTCGG
ACATCGGCCTGTATGGGTGCTGGTTCAAGTCCCAGATTTACGATGAGGAGGCCACCTGGGAGCTG
CGGGTGGCAGCACTGGGCTCACTTCCTCTCATTTCCATCGTGGGATATGTTGACGGAGGTATCCA
GTTACTCTGCCTGTCTCAGGCTGGTTCCCCAGCCACAGCCAAGTGGAAAGGTCCACAAGGAC
AGGATTTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC
ATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTCCATCCACCTTGCTGAGCAGAGTCATGA
GGTGAATCCAAGGTATTGATAGGAGAGACGTTTTTCCAGCCCTCACCTTGGCGCCTGGCTTCTA
TTTTACTCGGGTTACTCTGTGGTGGCTGTGTGGTGTGTGTCATGGGGATGATAATTGTTTTCTTC
AAATCCAAAGGAAAATCCAGGCGGAAGTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAG
AGACGCCCCGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCG
TTTCTGATCTGAAAAGTGAACCCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGA
TTTACAAGGAAGAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTGCGGATGACGTAGACAGGGGGAAGAACA
ATGTGACTTTTGTCTCCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTTGTATTTT
ACATTCAATCCCCATTTTATCAGCCTCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCCT
GGACTATGAGGGTGGGACCATCTCCTTCTTCAATACAAATGACCAGTCCCTTATTTATACCCTGC
TGACATGTCAAGTTGAAGGCTTGTGAGACCCTATATCCAGCATGCGATGTATGACGAGGAAAAG
GGGACTCCCATATTCATATGTCCAGTGTCTGGGGATGAGACAGAGAAGACCCTGCTTAAAGGGC
CCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGTCTCCGACAGGTGGCCCCAGCTTCTCTCT
CCGGAGCCTGCGCACAGAGAGTCACGCCCCCACTCTCCTTTAGGGAGCTGAGGTTCTTCTGCCC
TGAGCCCTGCAGCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGATTGGCCTGACCTGTGGGAG
TCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTTAGTTTGTGAAAA
CTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCCCAGGCTCCTCATTTGCTAGTCACGG
ACAGTGATTCTGCCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGCAGGTT
TGAGGGCACAGTGTGTTGCTAATGATGTGTTTTTATATTATACATTTCCACCATAAACTCTGTT
TGCTTATTCCACATTAATTTACTTTTCTCTATACCAATCACCCATGGAATAGTTATTGAACACC
TGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTCACTGATTCTATAAGCCCAGCAT
TACCTGATACCAAACCAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCATATCC
CTCATTAACACAGACACAAAATTTCTAAATAAAATTTTAACAAATTAACATAAATAATATTTA
AAGATGATATATAACTACTCAGTGTGGTTGTCCCACAAATGCAGAGTTGGTTTAAATTTAAAT
ATCAACCAGTGTAATTCAGCACATTAATAAAGTAAAAAGAAAACCATAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH
LYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGQDLSSDSRANADGYSLYDVEISI
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALCGVVMGMIIVFFK
SKGKIQAELDWRRKHGQAELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAPOEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDRGKNNVTLSPNNGYWVLRLTTEHLYFT
FNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLTLCQFEGLLRPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTAGGAAAAGAG
TTTGTGGGAACCCCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTCTGGCAGTGTGCATTGGA
CTCACTGTTTATTATGTGAGATATAATCAAAGAAGACCTACAATTACTATAGCACATTGTCAATTTACAAC
TGACAAACTATATGCTGAGTTTGGCAGAGAGGCTTCTAACAATTTTACAGAAATGAGCCAGAGACTTGAAT
CAATGGTGAAAAATGCATTTTATAAATCTCCATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTC
AGTCAACAGAAGCATGGAGTGTGGCTCATATGCTGTTGATTGTAGATTTCACTCTACTGAGGATCCTGA
AACTGTAGATAAAATTGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAG
ATCCTCACTCAGTTAAATTAATAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGGA
ACACGAAGAAGTAAACTCTAGGTCAGAGTCTCAGGATCGTTGGTGGGACAGAAAGTAGAAGAGGGTGAATG
GCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGTGGAGCAACCTTAATTAATGCCACATGGC
TTGTGAGTGTGCTCACTGTTTTACAACATATAAGAACCCTGCCAGATGGACTGCTTCCTTTGGAGTAACA
ATAAAACCTTCGAAAATGAAACGGGGTCTCCGGAGAATAATTGTCCATGAAAAATACAAACACCCATCACA
TGACTATGATATTTCTCTTGACAGACTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTG
TCCCTGATGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTTGTGACAGGATTGGAGCACTGAAAAAT
GATGTTTACAGTCAAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTGCAATGAACC
TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTTAGAAGGAAAAACAGATGCAT
GCCAGGTGACTCTGGAGGACCACTGGTTAGTTCAGATGCTAGAGATATCTGGTACCTTGCTGGAATAGTG
AGCTGGGAGATGAATGTGCGAAACCCAAACAGCCTGGTGTGTTTACTAGAGTTACGGCCTTGCGGACTG
GATTACTTCAAAAACCTGGTATCTAAGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTTGTTTTTG
GGTGTGGAGGCCATTTTTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTTGACTGATCTCA
ATAAACTGTTTGCTTGATGCATGTATTTCTTCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTGCCA
GATCAACTCTGTCTGTGAGCAATAGTTGAACTTTATGTACATAGAGAAATAGATAATACAATATTAC
ATTACAGCCTGTATTCATTGTTCTCTAGAAGTTTGTGAGAATTTGACTTGTGACATAAATTTGTAAT
GCATATATACAATTTGAAGCACTCCTTTTCTTCAGTTCCCTCAGCTCCTCTCATTTCAAGCAATATCCATTT
TCAAGGTGCAGAACAGGAGTGAAGAAAAATATAAGAAGAAAAAATCCCTACATTTTATTGGCACAGAA
AAGTATTAGGTGTTTTTCTTAGTGAATATTAGAAATGATCATATTCATTATGAAAGGTCAAGCAAAGACA
GCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGAACTAAGTTAAGGAAGTCCAGAAAGAAGCCAAG
ATATATCCTTATTTTCAATTTCCAAACAATACTATGATAAATGTGAAGAAGATTCTGTTTTTTGTGACCT
ATAATAATTATACAACTTCATGCAATGTACTTGTCTAAGCAAATTAAGCAAATATTTATTTAACATTG
TTACTGAGGATGTCAACATATAACAATAAATATAAATCACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYSTLSFTTDKLY
AEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKRGL
RRIIVHEKYKHPSHDYDISLAELSSPVPTNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDSSGGLVSSDARDIWYLAG
IVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:

amino acids 21-40 (type II)

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMDYLNNVLTLTAE TRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLLLLLAWYFGSLLVIFCVELACGVWVTYEQELMVPVQWSDMVT
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTMDWPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGCGKMYSLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDRREPGTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACCTCATCCTCATCCTCTTCTCTGATAAAGCCCCCTACCAGTGCT
GATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCTTAAAAAAAAGTGCTTGAAAGAGAAGGGGACAAAGGAACA
CCAGTATTAAGAGGATTTTCCAGTGTTTCTGGCAGTTGGTCCAGAAGGATGCTCCATTCTCTGCTTCTCACCTG
CCTCTTCATCACAGGCACCTCCGTGTCAACCCGTGGCCCTAGATCCTTGTCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCAGTTGGATGAGTCTCAAGGTCTCTCTATGTGACAACCATGTGAATGGGGAG
TGGTACCACCTTCACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAACCCA
CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAGGCACGGCATTGTGCAACGCCAGGCTTGTGCCAGCT
TCAATGGGAACTGCTGTCTCTGGAACACCACGGTGGAGTCAAGGCTTGGCCCTGGAGGCTACTATGTGTATCGT
CTGACCAAGCCCAGCGTCTGCTTCCACGTCTACTGTGGTCATTTTTATGACATCTGCGACGAGGACTGCCATGG
CAGCTGCTCAGATACCAGCGAGTGCACATGCCTCCAGGAAGTGTGCTAGGCCCTGACAGGCAGACATGCTTTG
ATGAAATGAATGTGAGCAAAACAACGGTGGCTGCAGTGAGATCTGTGTGAACCTCAAAAACCTCCTACCGCTGT
GAGTGTGGGGTGGCCGTGTGCTAAGAAGTGATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA
TGGTGGCTGCAGCCACTCTTGCCTTGGATCTGAGAAAGGCTACCAGTGTGAATGTCCCCGGGGCCTGGTCTGT
CTGAGGATAACCACACTTGCCAAGTCCCTGTGTGTGCAAAATCAAATGCCATTGAAGTGAACATCCCCAGGGAG
CTGGTTGGTGGCCTGGAGCTCTTCTGACCAACACCTCCTGCCGAGGAGTGTCCAACGGCACCCATGTCAACAT
CCTCTTCTCTCTCAAGACATGTGGTACAGTGGTCGATGTGGTGAATGACAAGATTGTGGCCAGCAACCTCGTGA
CAGGTCTACCAAGCAGACCCCGGGGAGCAGCGGGGACTTCATCATCCGAACCAGCAAGCTGCTGATCCCGGTG
ACCTGCGAGTTTCCACGCCGTGTACACCAATTTCTGAAGGATACGTTCCCAACCTTCGAAACTCCCCACTGGAAAT
CATGAGCCGAATCATGGGATCTTCCCATTCACTCTGGAGATCTTCAAGGACAATGAGTTTGAAGAGCCTTACC
GGGAAGCTCTGCCCACCCTCAAGCTTCGTGACTCCCTCTACTTTGGCATTGAGCCCGTGGTGCACGTGAGCGGC
TTGGAAGCTTGGTGGAGAGCTGCTTTGCCACCCACCTCCAAGATCGACGAGGTCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTTCAGATGACTCGGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTTCCAGG
TCCCTGTCTTCAAGTTTGTGGGCAAAGACCACAAGGAAGTGTCTGCACTGCCGGGTTCTTGTCTGTGGAGTG
TTGGACGAGCGTTCCCGCTGTGCCAGGGTTGCCACCGCGCAATGCGTCGTGGGGCAGGAGGAGGACTCAGC
CGGTCTACAGGGCCAGACGCTAACAGCGGCGCCGATCCGCATCGACTGGGAGGACTAGTTTCGTAGCCATACCTC
GAGTCCCTGCATTGGACGGCTCTGCTCTTTGGAGCTTCTCCCCCACC GCCCTCTAAGAACATCTGCCAACAGC
TGGGTTTCAGACTTCACACTGTGAGTTCAGACTCCAGCACCAACTCACTCTGATTCTGGTCCATTCACTGGGCA
CAGGTACAGCACTGCTGAACAATGTGGCTGGGTGGGTTTCATCTTTCTAGGGTTGAAAACCTAACTGTCCA
CCAGAAAGCACTCACCCATTTCCTCATTTCTTCTACACTTAAATACCTCGTGTATGGTGCAATCAGAC
CACAAAATCAGAAGCTGGGTATAATATTCAAGTTACAAACCTAGAAAAATTAACAGTTACTGAAATTATGA
CTTAAATACCAATGACTCCTTAAATATGTAAATTATAGTTATACCTTGAATTTCAATTCAAATGCAGACTAA
TTATAGGGAATTTGGAAGTGTATCAATAAAACAGTATATAATTTT

FIGURE 110

MPPFLLLTCLFITGTSVSPVALDPCSAYISLNEPWNRNTDHLQDESQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPGGYVYRLTKPSVCFHV
YCGHFYDIDCEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV
LRSDGKTCEDVEGCHNNNGGCSHSLGSEKGYQCECPRGLVLSEDNHTCQVPVLCKSNAIEVNIIPRELVGG
LELFLTNTSCRGVSNNGTHVNILFSLKTCGTVDVVDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEOPYREALPTLKLKRDLSLYFGIEPVVHV
SGLESLVESCFATPTSKIDEVLKYYLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDHKEVFLHCRV
LVCVGLDERSRCAQGGCHRRMRMRGAGGEDSAGLQGQTLTGGPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG
GCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGGAGGCCCT
CCTGCTGCCCTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCCGTATCCCC
ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTTGT
GGTTGTCTCATCAAGGTGATTCTGGATAAATACTACTTCTCTGCGGGCAGCCTCTCCACTTCATCCCGA
GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
CCCGAAGGGCCTGCAGTGGCAGTCCGCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
AGGGAAGTGGTTCTCTGCTGTTTCGACAATTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATGG
GCTACAGCAGAGCTGTGGAGATTGGCCAGACCAGGATCTGGATGTTGTTGAAATCACAGAAAACAGCCAG
GAGCTTCGCATGCGGAACCTCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCAGTGTCTTGCCCTG
TGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGTGGGGAGGAGGCCCTCTGTGGATTCTTGCCCTTGGCAGG
TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCTCACGGCA
GCCCCTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAAATGGGGCAG
CTTCCCATCCCTGGCTGTGGCCAAGATCATCATATTGAATTCAACCCATGTACCCCAAAGACAATGACA
TCGCCCTCATGAAGCTGCAGTTCCTCACTTTCTCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTT
GATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGATGGGGCTTTACGAAGCAGAATGGAGGGAA
GATGTCTGACATACTGCTGCAGGCGTCAGTCCAGGTCATTGACAGCACACGGTGCAATGCAGACGATGCCGT
ACCAGGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGGAAGGGGGTGTGGACACCTGCCAGGGT
GACAGTGGTGGGCCCCCTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGG
CTGCGGGGGCCCCGAGCACCCCAGGAGTATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT
GGAAGGCTGAGCTGTAATGTCTGCTGCCCTTTGCAGTGTGGGAGCCGCTTCCTTCCTGCCCTGCCACCT
GGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCCCTGGGTACACCCCTCTGCCACAGCCTCAGCAT
TTCTTGAGCAGCAAAGGGCCTCAATTCCTGTAAGAGACCCCTCGCAGCCAGAGGCGCCAGAGGAAGTCA
GCAGCCCTAGCTCGGCCCACTTGGTGTCTCCAGCATCCAGGGAGAGACACAGCCCACTGAACAAGGTCT
CAGGGGTATTGCTAAGCCAAGAAGGAACTTCCACACTACTGAATGGAAGCAGGCTGTCTTGTAAGGCC
CAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGCTCTGCGCCAGCCCTGTCCGTCTTACCCATCCCCAA
GCCTACTAGAGCAAGAAACCAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT
ACTGTTGTCAATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAA
AAAA

FIGURE 112

MLQDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPIIIIALLSLASIIIVVLIKVILDKYYFLCG
QPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN
FTEALAEACRQMGYSRAVEIGPDQDLDVVEITENSQELMRNSSGPCLSGSLVSLHCLACGKSL
KTPRVVGGEASVDSWPWQVSIQYDKQHVCSSILDPHWLTAHCFRKHTDVFNWKVRAGSDKL
GSFPSLAVAKIIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTQCQDSSGGPLMYQS
DQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCTGGTCCCAAGTGATCCACCCGCTCAGCCTCCCAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACCTTTTTCAGCAACTAAAAAGCCACAGGAGT
TGAAC TGCTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTACATTAAAAATC
TGTTTTTTGTTCTCTTGTAACTAGCCTTTACCTTCCTAACACAGAGGATCTGTCACTGTGGCTCT
GGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTCTACCCACACCGTCCCCTCGAAG
CCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTCACCAACTGTCTCACGTCT
GGAGGCACTGACTCGGGCAGTGCAAGTAGCTGAGCCTCTTGGTAGCTGCGGCTTTCAAGGTGGGC
CTTGCCCTGGCCGTAGAAGGGATTGACAAGCCCCGAAGATTTTCATAGGCGATGGCTCCCACTGCCC
AGGCATCAGCCTTGCTGTAGTCAATCACTGCCCCGGGGCCAGGACGGGCCGTGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACGCTGCCCCGCCATCTAACCTTTTCATGTCCTGCACATCACCTG
ATCCATGGGCTAATCTGAACTCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGGTCTGCTTAGACCACCTGGTTTATGTGACAGGACTTGCATTCTCCTGGAACATGAGG
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAACTTGTGCCAAATTATGGGTGAGAAAAGATG
GAGGTGTTGGGTTATCACAAGGCATCGAGTCTCCTGCATTGAGTGACATGTGGGGGAAGGGCTG
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCCGAT
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAA
TCTGCGATCACCAGCCAGGGGCAGCCGCTTGGGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCC
CTCCTTCCCTCTGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTTGGCCAGGAGGTGAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCT
GTGCGAACACGGAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCAGGCAAAGGACTGTGTGG
CTCAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCAAGACCAAAGGAGCTAGAGCTTGGTT
CAAATGATCTCCAAGGGCCCTTATACCCAGGAGACTTTGATTTGAATTTGAAACCCCAAATCCA
AACCTAAGAACCAGGTGCATTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACAT
TTTGGGAGGCCGAGGCGGGTAGATCACCTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAACATGG
TGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACTGGGAGGTGAAGGAGGCTGAGACA
GGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAAATAAAAAAGAATTA
TGTTTATTTGTAA

FIGURE 114

MLWWLVLLLLPTLKSVCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFGQGPCPGRRRD

Signal peptide:
amino acids 1-15

FIGURE 115

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAAGAAA
ATATGTAAATCACTTAAGATTTGTGGACTGGTGTGTTGGTATCCTGGCCCTAACTCTAATTGTCCCT
GTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGATTTACATGGAAATTGATCCTGTGACCAGAACTGAAATA
TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATACACTGG
CATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGTGATTCTGAATTTT
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACCTTTCTTTGAACAGTCAGTG
ATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATTCCAAAATTCTGGA
GATTTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAG
ACTTTGAGGAGGAGGGAGAAGATCTTCACTTTCTGCCAACGAAAAAAAGGGATTGAACAAAAT
GAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGA
AGAACTTCCAATAAATGACTATACTGAAAATGGAATAGAATTTGATCCCATGCTGGATGAGAGAG
GTTATTGTTGTATTTACTGCCGTCGAGGCAACCGCTATTGCCGCCGCGTCTGTGAACCTTTACTA
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCCCTTG
TAACTGGTGGGTGGCCCGCATGCTGGGGAGGGTCTAATAGGAGGTTTGAGCTCAAATGCTTAAAC
TGCTGGCAACATATAATAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCT
GGTAGCCAGCTCTCCAGAATTACTTGTAGGTAATTCCTCTCTTCATGTTCTAATAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFKSKICKSLKICGLVFGILALTILVLFWGSKHFWPEVPKKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLVHDFKNGYTGIIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEEITTTFFEQSVIWVPAEKPIENRDFLKNKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPANEKKGIEQNEQWVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMLD
ERGYCCIIYCRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGCAGGCGGCA
GGGCGGGCGGCCAGGATCATGTCCACCACCACATGCCAAGTGGTGGCGTTCCCTCCTGTCCATCCTGGGGCT
GGCCGGCTGCATCGCGGCCACCGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCCGTCACCT
CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTACCGAATGCAGGCCC
TATTTACCATCCTGGGACTTCCAGCCATGCTGCAGGCAGTGCAGGCCCTGATGATCGTAGGCATCGTCCT
GGGTGCCATTGGCCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTG
CCAAAGCCAACATGACACTGACCTCCGGGATCATGTTTCATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTG
TCTGTGTTTGCCAACATGCTGGTGACTTAACCTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGG
GATGGTGACAGTGTTCAGACCAGGTACACATTTGGTGCGGCTCTGTTCTGGGCTGGGTGCGTGGAGGCC
TCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGCGCTGGCACCAGAAGAAACCAACTACAAA
GCCGTTTCTTATCATGCCTCAGGCCACAGTGTTCCTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTT
TGGGTCCAACACCAAAACAAGAAGATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATC
CTTCCAAGCAGCACTATGTGTAATGCTCTAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCA
CCCAAAAACAAGGAGATCCCCTCTAGATTTCTTCTTGTCTTTGACTCACAGCTGGAAGTTAGAAAAGCCT
CGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCCACCATAAAACA
GCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTCTTTTTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCACATTTTGATGATTAGACAGACTCCCCCTC
TTCCTCCTAGTCAATAAAACCAATTGATGATCTATTTCCAGCTTATCCCAAGAAAACCTTTGAAAGGAAA
GAGTAGACCCAAAGATGTTATTTCTGCTGTTGAATTTGTCTCCCCACCCCAACTTGGCTAGTAATAA
ACACTTACTGAAGAAGAAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTT
ACACTGTGATCTTAAAGTTACCAACCAAGTCATTTTCAGTTTGAGGCAACCAACCTTTCTACTGCTG
TTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCCTCTTCTGT
CGCGGGTCAGAAATGTCCCTAGATGAATGAGAAAATTATTTTTTTTAAATTTAAGTCCTAAATATAGTTAA
AATAAATAATGTTTTAGTAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAGAAG
GAAATGAAAAATAATTGCTTTGACATTGTCTATATGGTACTTTGTAAAGTCATGCTTAAGTACAAATTC
ATGAAAAGCTCACACCTGTAATCTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAGT
TCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAATACAGAGAGAAAAATCAGCCAGTCA
TGGTGGCATAACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGGAGGT
TGGGGCTGCAGTGAGCCATGATCACACCCTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA
AATAAAAAATAAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTTAAACTAATTCCTTAA

FIGURE 118

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAACTGTTCTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCCGGAGTCC
AGCTGGCTAAAACTCATCCCAGAGGATAATGGCAACCCATGCCTTAGAAATCGCTGGGCTGTTTCTTG
GTGGTGTGGAATGGTGGGCACAGTGGCTGTCACTGTCATGCCTCAGTGGAGAGTGTGGCCTTCATT
GAAAACAACATCGTGGTTTTTGAAAACCTTCTGGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
CATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTTTCTCCGGACCTACAGGCAGCCAGAG
GACTGATGTGTGCTGCTTCCGTGATGTCTTCTTGGCTTTCATGATGGCCATCCTTGGCATGAAATGC
ACCAGGTGCACGGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCAT
CATCACGGGCATGGTGGTGTCTCATCCCTGTGAGCTGGGTTGCCAATGCCATCATCAGAGATTTCTATA
ACTCAATAGTGAATGTTGCCCAAAAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCA
CTGGTGTGATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTTGTTGCAACGAAAAGAGCAGTAGCTA
CAGATACTCGATACCTTCCCATCGCACAAACCAAAAAGTTATCACACCGGAAAGAAGTCAACGAGCG
TCTACTCCAGAAGTCAGTATGTGTAGTTGTGTATGTTTTTTAACTTTACTATAAAGCCATGCAAATG
ACAAAAATCTATATTACTTTCTCAAAATGGACCCCAAGAACTTTGATTTACTGTTCTTAACTGCCT
AATCTTAATTACAGGAAGTGTGCATCAGCTATTTATGATTCTATAAGCTATTTACGAGAATGAGATA
TTAAACCAATGCTTTGATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAAGCATCTA
CTCTTTTTATCATTACTTCAAAATGACATTGCTAAAGACTGCATTATTTTACTACTGTAATTTCTCC
ACGACATAGCATTATGTACATAGATGAGTGTAAACATTTATATCTCACATAGAGACATGCTTATATGGT
TTTATTTAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCACTATTGCTTTTCAGGGAA
ATCATGGATAGGGTTGAAGAAGGTACTATTAATTGTTTAAAAACAGCTTAGGGATTAATGTCCTCCA
TTTATAATGAAGATTTAAATGAAGGCTTTAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATAT
GCTGTTTTTTAGCCTAGGAGTTAGAAATCCTAAGTTCTTTATCCTCTTCTCCAGAGGCTTTTTTTTT
CTTGTGTATTAAATTAACATTTTAAAACGCAGATATTTTGTCAAGGGGCTTTCATTCAAACCTGCTT
TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATGGTTTTAGGAAAGTG
AAAATATTTTGTTTTTGTATTTGAAGAAGAATGATGCATTTTGACAAGAAATCATATATGTATGGAT
ATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAAGAGCAGAAAAATA
TGCTTGGTTTTTCATTTGCTTACCAAAAAACAACAAAAAAGTTGTCCTTTGAGAACTTCACCT
GCTCCTATGTGGGTACCTGAGTCAAAATTGTCATTTTGTCTGTGAAAAATAAATTCCTTCTTGTA
CCATTTCTGTTTAGTTTTACTAAAACTGTAAATACTGTATTTTCTGTTTATCCAAATTTGATGAA
ACTGACAATCCAATTTGAAAGTTTGTGTCGACGTCTGTCTAGCTTAAATGAATGTGTTCTATTGCTT
TATACATTTATATTAATAAATTGTACATTTTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTVAVTVMQWRVSAFIENNIVVFENFW EGLWMNCVRQANIRMQCK
IYDSLLALSPDLQAARGLMCAASVMSFLAFMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
MVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCNEKSSSY
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCGGAG
CCAGACGCTGACCACGTTCTCTCCTCGGTCTCCTCCGCTCCAGCTCCGCGCTGCCCCGGCAGCC
GGGAGCCATGCGACCCCAGGGCCCCGCCGCTCCCCGCAGCGGCTCCGCGGCCTCCTGCTGCTCC
TGCTGCTGCAGCTGCCCCGCGCCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTGGGATCCCAGGTCCGGATG
GATTCAAAGGAGAAAAGGGGAATGTCTGAGGGAAAGCTTTGAGGAGTCCTGGACACCCAACTAC
AAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCTTGGGAAAATTGCGGAGTGTACATT
TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGTTTCAGTGGCTCACTTCGGCTAAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTTACATTCAATGGAGCTGAATGTTTCAGGACCTCTTCCC
ATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAATATTCATCG
CACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGCTGGATTAGTGGATGTTGCTATCTGGG
TTGGCACTTGTTTCAGATTACCCAAAAGGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATC
ATTATTGAAGAACTACCAAAATTAATGCTTTAATTTTCATTGCTACCTCTTTTTTTATTATGCC
TTGGAATGGTTCACCTAAATGACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAG
CTAAATATGTTTACAGACCAAAGTGTGATTTACACTGTTTTTAAATCTAGCATTATTCATTTTG
CTTCAATCAAAAGTGTTTCAATATTTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATT
CTCTCAACCTATAATTTGGAATATTGTTGTGGTCTTTTGTTTTTTCTTTAGTATAGCATTTTTA
AAAAAATATAAAAGCTACCAATCTTTGTACAATTTGTAAATGTTAAGAATTTTTTTTATATCTGT
TAAATAAAAATTATTTCCAACA

FIGURE 122

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAQVPGR
DGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECGSLPIEAIYLDQGSPEMNSTINIHRTS
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTGAA
 CTGGGTGCTCATCACGGGAACCTGCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
 ATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTTTTTTA
 ACCGCCCCCTCCCCACCCCCCAAAAAAAGCTGTAAGATGCAAAAACGTAATATCCATGAAGATCC
 TATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTTGTTCTTGGAG
 TGTTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCCAAGGGGTCCAATTTT
 TCTTCCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGGCTGTCATGCAACTG
 GCCCCAAGCCAAAGCAAAAGACCTAAGGACGACCTTTGAACAATACAAAGGATGGGTTTCAATG
 TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCACTGTCTTACTGACAATG
 CTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAAGGCAAAATGGTATATTGTGA
 ATCTCAGAAATTACAGGAGATACCCCTCAAGTATATCTGCTGGTTGCTTAGGTTTGTCCCTTCGCT
 ATAACAGCCTTCAAAAACCTAAGTATAATCAATTTAAAGGGCTCAACCAGCTCACCTGGCTATAC
 CTTGACCATAACCATATCAGCAATATTGACGAAAATGCTTTTAAATGAATACGCAGACTCAAAGA
 GCTGATTCTTAGTTCCAATAGAATCTCTATTTTCTTAACAATACCTTCAGACCTGTGACAAATT
 TACGGAACCTGGATCTGTCTATAATCAGCTGCATTCTCTGGGATCTGAACAGTTTCGGGGCTTG
 CGGAAGCTGCTGAGTTTACATTTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCA
 AGACTGCCGCAACCTGGAACCTTTGGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATG
 TCTTTGCTGGCATGATCAGACTCAAAGAACTTCACCTGGAGCACAAATCAATTTTCCAAGCTCAAC
 CTGGCCCTTTTCCAAGGTTGGTCAGCCTTCAGAACCTTTACTTGCAGTGGAATAAAATCAGTGT
 CATAGGACAGACCATGTCTGGACCTGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGA
 TCGAAGCTTTTCAGTGGACCCAGTGTTTTCCAGTGTGTCCCGAATCTGCAGCGCCTCAACCTGGAT
 TCCAACAAGCTCACATTTATTTGGTCAAGAGATTTTGGATTCTTGGATATCCCTCAATGACATCAG
 TCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTGCTCCCTTGTAACCTGGCTGAAAAGTT
 TTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTG
 ATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAAGTACTACAGAGAGGTTTGATCTGGCCAG
 GGCTCTCCCAAAGCCGACGTTTAAGCCCAAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCTT
 TGCCCCGACGGTGGGAGCCACAGAGCCCGGCCAGAGACCGATGCTGACGCCGAGCACATCTCT
 TTCCATAAAATCATCGCGGGCAGCGTGGCGCTTTTCTGTCCGTGCTCGTCATCCTGCTGGTTAT
 CTACGTGTCATGGAAGCGGTACCCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAA
 GGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCCAGCACCCAGGAATTTTATGTA
 GATTATAAACCCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCTGCACCTA
 TAACAAATCGGGCTCCAGGGAGTGTGAGGTATGAACCATTTGTGATAAAAAGAGCTCTTAAAGCT
 GGGAAATAAGTGGTGCTTTTATGAACTCTGGTGACTATCAAGGGAACGCGATGCCCCCCTCCCC
 TTCCCTCTCCCTCTCACTTTTGGTGGCAAGATCCTTCCTTGTCCGTTTTAGTGCATTCAATACT
 GGTCATTTTCTCTCATACATAATCAACCCATTGAAATTTAAATACCACAATCAATGTGAAGCTT
 GAACTCCGGTTTAAATATAATACCTATTGTATAAGACCCTTTACTGATTCCATTAATGTGCAATTT
 GTTTTAAGATAAAACTTCTTTCATAGGTAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGCRCEGKMVYCESQKLQEIPSSISAGCLG
LSLRYNLSLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR
PVTNLRNLDSYNQLHSLGSEQFRGLRKLSSLHLRSNSLRTIPVRI FQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKEHLEHNQFSKLNALFPRLVSLQONLYLQWNKISVIGQTMSWTWSSLQRLDL
SGNEIEAFSGPSVVFQCPNLQRLNLDNKLTFIGQEILDSWISLNDISLAGNIWECSRNICSLVN
WLKSFKGLRENTIICASPKELOQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMKQLQOR
SIMRRHRKKKRQSLKQMPSTQEFYVDYKPTNTTETSEMLLNGTGPCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCGCTACTGCTGAATGTCGTCGCCGAGGAGGAGGAGAGGCTTTTGCCGCTG
ACCCAGAGATGGCCCCGAGCGAGCAAATTCCTACTGTCCGGCTGCGGGCTACCGTGGCCGAGCT
AGCAACCTTTCCCCTGGATCTCACAAAACTCGACTCCAAATGCAAGGAGAAGCAGCTCTTGCTC
GGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCATTTACAGACACGTAGT
GTATTCTGGAGGTGCAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTTGGCAAAAGTGAAG
ATGAGCATTATCCCCTTTGAAATCAGTCATTGGAGGGATGATGGCTGGTGTATTGGCCAGTTT
TTAGCCAATCCAAGTACCTAGTGAAGGTTGAGATGCAAATGGAAGGAAAAAGGAACTGGAAGG
AAAACCATTGCGATTTCTGTTGTACATCATGCATTTGCAAAAATCTTAGCTGAAGGAGGAATAC
GAGGGCTTTGGGCAGGCTGGGTACCCAATATACAAAGAGCAGCACTGGTGAATATGGGAGATTTA
ACCACCTTATGATACAGTGAAACACTACTTGGTATTGAATACACCACTTGAGGACAATATCATGAC
TCACGGTTTATCAAGTTTATGTTCTGGACTGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCA
TCAAAGCAGAATAATGAATCAACCACGAGATAAACAAGGAAGGGGACTTTTGTATAAATCATCG
ACTGACTGCTTGATTGAGCTGTTCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTTTACC
ATCTTGGCTGAGAATGACCCCTGGTCAATGGTGTCTGGCTTACTTATGAAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTTTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGARES
APYRGMVRTALGIIEEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLEVVFGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMMEGKRRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
NIQRAALVMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTPADVIKSRIMNQP
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGARES
APYRGMVRTALGIIEEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLEVVFGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMMEGKRRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
NIQRAALVMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTPADVIKSRIMNQP
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

FIGURE 127

CGCGGATCGGACCCAAGCAGGTCGGCGGCGGGCGGCAGGAGAGCGGCCGGGCGTCAGCTCCTCGAC
CCCCGTGTCGGGCTAGTCCAGCGAGGCGGACGGGCGGCGTGGGCCCATGGCCAGGCCCGGCATGG
AGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCGCTCGGGGGGCATCGGCGCGGGCGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGACTTTGATCCCCCTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC
TGCATCAACAATGCTGGCTTGGCCCGGCGCTGACACCCTGCTCTCAGGCAGCACCAGTGGTTGGAA
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA
AGGAGCGGAATGTGGACGATGGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTTA
CCCCGTGCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGGACT
GAGGCAAGAGCTTCGGGAGGCCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGG
AGACACAATTGCCTTCAAACCTCCACGACAAGGACCCTGAGAAGGCAGCTGCCACCTATGAGCAA
ATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCTCAGCACCCCCGCACA
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTAGTGACTGTGGGAGCTCC
TCCTTCCCTCCCCACCCTTCATGGCTTGCCCTCCTGCCTCTGGATTTTAGGTGTTGATTTCTGGAT
CACGGGATACCACTTCCTGTCCACACCCCGACCAGGGGCTAGAAAATTTGTTTGAGATTTTATA
TCATCTTGTCAAATTGCTTCAGTTGTAATGTGAAAAATGGGCTGGGGAAAGGAGGTGGTGTCCC
TAATTGTTTTACTTGTTAACTGTTCTTGCGCCCTGGGCACTTGGCCTTTGTCTGCTCTCAGTG
TCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATCCCCATCTTCTGCACCTCAACGTCTG
TGGCTCAGGGCTGGGGTGGCAGAGGGAGGCCTTACCTTATATCTGTGTTGTTATCCAGGGCTCC
AGACTTCCTCCTCTGCCTGCCCCACTGCACCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCC
AGCCCAGTCTTGGCTTCTTGTCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAG
CAGAACACCAGGCCTGGCCCACTGGATTTCATGGTGATCATTAAAAAAGAAAAATCGCAACCAA
AAAAAAAAA

FIGURE 128

MARPGMERWRDRLALVTGASGGIGAAVARALVQQGLKVVGCAITVGNIEELAAECKSAGYPGTLI
PYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPDTLTSGSTSGWKDMFNVNVLALSICTR
EAYQSMKERNVDDGHIININMSGHRVPLSVTHFYATKYAVTALTEGLRQELREAQTHIRATC
ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTTCTACATGGGCCTCCTGCTGCTGGTGTCTTCCTCAGCCTCCTGCCGGTGGCCTACACCAT
CATGTCCCTCCCACCCTCCTTTGACTGCGGGCCGTTCAAGTGCAGAGTCTCAGTTGCCCGGGAGC
ACCTCCCCTCCCGAGGCAGTCTGCTCAGAGGGCCTCGGCCCAGAATTCCAGTTCTGGTTTCATGC
CAGCCTGTAAAAGGCCATGGAACCTTGGGTGAATCACCGATGCCATTTAAGAGGGTTTTCTGCCA
GGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTTTCATTTCACTAGCCACCAGCCACCTGTGG
CCGTTGAGTGCTTGAAATGAGGAAGTGAAGAAATTAATTTCTCATGTATTTTTCTCATTTATTTA
TTAATTTTTAACTGATAGTTGTACATATTTGGGGGTACATGTGATATTTGGATACATGTATACAA
TATATAATGATCAAATCAGGTAAGTGGGATATCCATCACATCAAACATTTATTTTTTATTCTTT
TTAGACAGAGTCTCACTCTGTACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCCAAGTAGCTGGGACTACAGGCAT
GCACCACAATGCCCAACTAATTTTTGTATTTTTAGTAGAGACGGGGTTTTGCCATGTTGCCCAGG
CTGGCCTTGAACTCCTGGCCTCAAACAATCCACTTGCCTCGGCCTCCCAAAGTGTATGATTACA
GGCGTGAGCCACCGTGCCTGGCCTAAACATTTATCTTTCTTTGTGTGGGAACTTTGAAATTAT
ACAATGAATTATTGTTAACTGTCTCTCCCTGCTGTGCTATGGAACACTGGGACTTCTTCCCTCT
ATCTAACTGTATATTTGTACCAGTTAACCAACCGTACTTCATCCCCACTCCTCTCTATCCTTCCC
AACCTCTGATCACCTCATTTCTACTCTCTACCTCCATGAGATCCACTTTTTTAGCTCCCATGTG
AGTAAGAAAATGCAATATTTGTCTTTCTGTGCCTGGCTTATTTCACTTAACATAATGACTTCCTG
TTCCATCCATGTTGCTGCAAATGACAGGATTCGTTCTTAATTTCAATTAAATAACACACATG
GCAAAAA

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRIPVLVSCQPV
KGGTILGESPMPPFKRVFCQDGNVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
ATCTTCCTCATCGGGACTAAAATTGGGCTGTTCCCTTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
TCCATCTGTGTGCTCGCTGCGATGCGGGTTTCATTTACTGTAATGATCGCTTCTGACATCCATTCCAACAG
GAATACCAGAGGATGCTACAACCTCTCTACCTTCAACAACCAAATAAATAATGCTGGGATTCTTCAGAT
TTGAAAACTTGCTGAAAGTAGAAAGAATATACCTATACCACAACAGTTTAGATGAATTCCTACCAACCT
CCCAAAGTATGTAAAAGAGTTACATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTTCAA
AAATTCCTATCTGGAAGAATTACATTTAGATGACAACCTCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCA
TTCCGAGACAGCAACTATCTCCGACTGCTTTTCTGTCCCGTAATCACCTTAGCACAATCCCTGGGGTTT
GCCCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTTTCATCACCATCTCTCAAG
GTCTCACTAGTCTAAAACGCCTGGTTCTAGATGGAAACCTGTTGAACAATCATGGTTTAGGTGACAAAGTT
TTCTTCAACCTAGTTAATTTGACAGAGCTGTCCCTGGTGCAGAAATCCCTGACTGCTGCACCAGTAAACCT
TCCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATAACCACATCAATCGGGTGCCCCAAATGCTTTTT
CTTATCTAAGGCAGCTCTATCGACTGGATATGTCCAATAATAACCTAAGTAATTTACCTCAGGGTATCTTT
GATGATTTGGACAATATAACACAACCTGATTCTTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATG
GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTCATGTGCCAAGCCCCAGAAA
AGGTTCTGTTGGGATGGCTATTAAGGATCTCAATGCAGAACCTGTTTGATTGTAAGGACAGTGGGATTGTAAGC
ACCATTGAGATAACCACTGCAATACCCAACACAGTGTATCTGCCCCAAGGACAGTGGCCAGCTCCAGTGAC
CAAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAAACCACAGGGAGTCCCTCAAGAAAAA
CAATTACAATTACTGTGAAGTCTGTACCTCTGATACCATTATCTCTTGGAACCTTGCTCTACCTATG
ACTGCTTTGAGACTCAGCTGGCTTAACTGGGCCATAGCCCGGCATTTGGATCTATAACAGAAACAATTGT
AACAGGGGAACGCAGTGAGTACTTGGTCACAGCCCTGGAGCCTGATTACCCCTATAAAGTATGCATGGTTC
CCATGGAACCAAGCAACCTCTACCTATTTGATGAAACTCCTGTTTGTATTGAGACTGAAACTGCACCCCTT
CGAATGTACAACCTACAACCACCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAACCCCAATTTACC
TTTGGCTGCCATCATTGGTGGGGCTGTGGCCCTGGTTACCATTGCCCTTCTTGCTTTAGTGTGTTGGTATG
TTCATAGGAATGGATCGCTCTTCTCAAGGAACGTGCATATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
GCAGAAGCTGGCACTAAGAAGGACAACCTATCTCTGAAATCAGGGAACTTCTTTTCAGATGTTACCAAT
AAGCAATGAACCCATCTCGAAGGAGGAGTTTGTAATACACACCATATTTCTCTTAATGGAATGAATCTGT
ACAAAAACAATCACAGTGAAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTTTGGGTTTTTTAAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRC DAGFIYCNDRFLT SIPTGIPEDATTLYL
QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTITYDSLSKIPYL
EELHLDNDSVSAVSIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
QGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRLQYRLDMSNNLSNLPQGI FDDL DNITQLILRNNPWYCGCKMKWVRDWLQSLPV
KVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAQQQWPAPVTKQPD
IKNPKLTKDQQTGSPSRKTTITVKSVTSDTIHISWKLALPMTALRLSWLKLGHSPAFGSITET
IVTGERSEYLVTALEPDSPYKVCMPMETSNLYLFDETPVC IETETAPLRMYNPPTTLNREQEKE
PYKNPNPLAAIIGGAVALVTIALLALVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKDNS
ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNLKNNHSESSNRSYRDSGIPDSHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCTTCCCAGAGTCCTTTGCCAGGCCACCCAGGCTTCTTGGA
GCCCTGCCGGGCCACTTGTCTTCATGTCTGCCAGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
TGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGGCCATGGCCAGCCTGGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACCGCTGTGGT
CCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT
GCGCTGGAGAGAGCCACCGTCTTCTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT
GGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCGGGAGAAGTGGGCCCAGGAGCCCCCTGC
TGCAGCCGCTGAGCCTGCGCGTGGGGATGCTGGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCCAGCTGACCCCTCCAGCCCGG
GTTTTTGAAGCTCCACATGCCTGGATCCACACTGATGCCTCCTTGGTGTACCCACGTTCTGGGC
CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCTGGTGCAGCTGCTGGGAACCGGGACG
GACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCCGGCTGCTC
AGGCTACTGCCTGTCCCACCACTGCTCTTCTCTCTGGGCCAGAATGAGGGGATGCACACAGG
GACCACTCCAACAGAGCCAGGACTATATCAACCTCTTCTGCGCCAACATGATGGACTTGAACCGC
AGAGCTGAGGCCATCGGATACGCCTACCCACCGGGACATCTTCATGGAAAACATCATGTTCTG
TGGAATGGGCGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
AACAGCAGGAAGGATGCTTCGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
TATCAGCAGCATTTTTTCGAGGAGAGTGAAGAGGCGAGAAAAACAATTTCCAGATTCTCGCTCTGT
TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTTGCCTCCTGGGTTCAAGC
AATTCTCTTGCTCATCCTCCCGAGTAGCTGGGACTACAGGAGCGTGCCACCATACTGGCTAAT
TTTTTATATTTTTTTAGTAGAGACAGGGTTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGAT
CTCAAGAGATCCGCCACCTCAGGCTCCCAAAGTGTTGGGATTATAGGTGTGAGCCACCGTGTCTG
GCTGAAAAGCACTTTCAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTC
ATGGGGGCTCTCTCCCTTAGATGGCTGCTCCTCCACAACACAGCCACAGCAGTGGCAGCCCTGG
GTGGCTTCTATACATCCTGGCAGAATACCCCCAGCAAACAGAGAGCCACCCCATCCACACCG
CCACCACCAAGCAGCCGCTGAGACGGACGGTTCCATGCCAGCTGCCTGGAGGAGGAACAGACCCC
TTTAGTCCTCATCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
ATAAGCAAAGCCACCCCGACACCCAATCTTGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAAAAAAA

FIGURE 134

MSARGRWEGGGRRACRGSGLGLARAQGAERTVSSEQRPAMASLGLLLLLLLLTALPPLWSSSLPGLD
TAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRSYLHKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGDSSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQSQD
YINLFCANMMDLNRRAEAIGYAYPTRDIEMENIMFCGMGGFSDFYKLRWLEAILSQKQEGCGF
EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFKQFSCILIP
SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCAATGGCGGCCGCTCTGTGGGGCTTCTTTCCCGTCCTGCTGCTG
CTGCTGCTATCGGGGGATGTCCAGAGCTCGGAGGTGCCCGGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGGCGTGCAAGTTGTTCCAGGGGTGAAGC
CTCAGGACTGGATCTCGGCGGCCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTTTCTTTAAG
ACAGATGGGAGTTTTGTGGTTCATGATATACCTTCTGGATCTTATGTAGTGGAAGTTGTATCTCC
AGCTTACAGATTTGATCCCGTTCGAGTGGATATCACTTCGAAAGGAAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCCAAATGAAATCTTCAGGT
CCACCTTCTTACTTTATTAAAAGGGAATCGTGGGGCTGGACAGACTTTCTAATGAACCAATGGT
TATGATGATGGTTCTTCCTTTATTGATATTTGTGCTTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT
GTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAATCATCTGGCAAATCTAGCAGCGGCAGCAG
TAAACAGGCAAAAGTGGGGCTGGCAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTGTCAC
AAACACGGCAACACTGGGTGGCATCCAAGTCTTGAAAACCGTGTGAAGCAACTACTATAAACTT
GAGTCATCCCGACGTTGATCTCTTACAACGTGTGTATGTT
AACTTTTGTAGCACATGTTTTGTACTTGGTACACGAGAAAACCCAGCTTTCATCTTTGTCTGTAT
GAGGTCAATATTGATGTCACTGAATTAATTACAGTGTCTATAGAAAATGCCATTAATAAATTAT
ATGAACTACTATACATTATGTATATTAATTAAACATCTTAATCCAGAAATCAAAAAAAAAAAAA
AAAAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLLSGDVQSSEVPGAAAEGSGSGVGIGDRFKIEGRAVVPGVKPDWISAA
RVLVDGEEHVGFLKTDGSEFVVHDIPSGSYVVEVVS PAYRFD PVRVDITSKGKMRARYVNYIKTSE
VVRLLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMVMMVLPLLI FVLLPKVVNTSDPDMRREME
QSMNMLNSNHELDPDVSEFMTRLFSSKSSGKSSSGSSKTGKSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGATTGATTCTCCCCAGTTCCTGTGGGTCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCAAACAA
GTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACCTCAGACCTGAAATCTTCTCTTACGGGAGGCTTGGCAGT
TTTTCTTACTCCTGTGGTCTCCAGATTTAGGCCTAAGATGAAAGCCTCTAGTCTTGCCTTCAGC
CTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAGACACTCAATTTGGG
AAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCTGAGATACGGGGCAGTG
TGCAAGCCAAAGATGGAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTTGCAAGACACA
AAGCCTGCGAATCGATGCTGCTCCTGCGCCATTTGCTAAGACTCTATCTGGACAGGGTATTTAA
AAACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGCAGCCTCGCCAATTCTTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCTCATGCCACATGACATGCCATTGTGGGGAGGAAGCAATG
AAGAAATACAGCCAGATTCTGAGTCACTTTGAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
TTTGGGGGAACTAGACATTCTTCTGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGATGCTG
CTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAA
ACCACCATCTCTTTACTGTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTG
CTTCCTTGATGATTGTCTTTATGCATCCCCAATCTTAATTGAGACCATACTTGTATAAGATTTT
TGTAATATCTTTCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTTTGCTATTTA
ATGTATTTATTTTTTACTTGGACATGAACTTTAAAAAAATTCACAGATTATATTTATAACCTG
ACTAGAGCAGGTGATGTATTTTTATACAGTAAAAAATAAACCTTGTAATCTAGAAAGAGTGG
CTAGGGGGGTATTTCATTGTATTCACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGAT
ATTTGAAATTGAACCAATGACTACTTAGGATGGGTTGTGGAATAAGTTTGTATGTGAATTGCAC
ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTTCCAG
CCAGGAATCTACACGGCCAGCATGTATTTCTACAAATAAAGTTTTCTTGCATACCAAAAAA
AAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWGQNGATTTCATNSHSDSELRPEIF
SSREAWQFFLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKIS
SLANSFLTIIKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTCGGTTCGCGATCCAGCC
TAGCGTGTCCACGATGCGGCTGGGCTCCGGGACTTTTCGCTACCTGTTGCGTAGCGATCGAGGTGC
TAGGGATCGCGGTCTTCCCTTCGGGGATTCTTCCCGGCTCCCGTTCGTTCTCTGCCAGAGCGGAA
CACGGAGCGGAGCCCCAGCGCCCGAACCCCTCGGCTGGAGCCAGTTCTAACTGGACCACGCTGCC
ACCACCTCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTTGAGAGATGATTTTGTGTTG
GGTCAAAGGGTGTGAAATTTATGCCCTACACAACCTTACCTTGTGGAAAAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGCATTGATGACGGGGAG
CCTTCTGGCTTTGTGCGACGTCATCAGGAACCTCAATTCTCCTGCACTGCTGGAAGACAGTGTGA
TAAGACAAGCAAAAGCAGCTGGAAAAAGAATAGTCTTTTATGGAGATGAAACCTGGGTTAAATTA
TTCCCAAAGCATTTTGTGGAATATGATGGAACAACCTCATTTTTCGTGTGAGATTACACAGAGGT
GGATAATAATGTCACGAGGCATTTGGATAAAGTATTAAGAGAGAGATTGGGACATATTAATCC
TCCACTACCTGGGGCTGGACACATTTGGCCACATTTTCAGGGCCCAACAGCCCCCTGATTGGGCAG
AAGCTGAGCGAGATGGACAGCGTGTGATGAAGATCCACACCTCACTGCAGTCGAAGGAGAGAGA
GACGCCCTTTACCCAATTTGTGGTTCTTTGTGGTGACCATGGCATGTCTGAAACAGGAAGTCACG
GGGCCTCCTCCACCGAGGAGGTGAATACACCTCTGATTTTAAATCAGTTCTGCGTTTGAAAGGAAA
CCCGGTGATATCCGACATCCAAAGCACGTCCAATAGACGGATGTGGCTGCGACACTGGCGATAGC
ACTTGGCTTACCGATTCCAAAGACAGTGTAGGGAGCCTCCTATTCCAGTTGTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTTACATTGGAATACAGTGCAGCTTAGTAAACTGTTGCAAGAG
AATGTGCCGTATATGAAAAAGATCCTGGGTTTGAGCAGTTTAAATGTGAGAAAGATTGCATGG
GAACTGGATCAGACTGTACTTGGAGGAAAAGCATTGAGAAGTCCTATTCAACCTGGGCTCCAAGG
TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTTGTCCCTGAGTGCACAAGTGGCCAG
TTCTCACCTGCTCCTGCTCAGCGTCCCACAGGCACCTGCACAGAAAGGCTGAGCTGGAAGTCCCA
CTGTCTATCTCCTGGGTTTCTCTGCTCTTTATTTGGTGATCCTGGTTCTTTCCGGCCGTTACGT
CATTGTGTGCACCTCAGCTGAAAGTTCGTGCTACTTCTGTGGCCTCTCGTGGCTGGCGGCAGGCT
GCCTTTTCGTTTACCAGACTCTGGTTGAACACCTGGTGTGTGCCAAGTGTGGCAGTGCCTGGAC
AGGGGGCCTCAGGGAAGGACGTGGAGCAGCCTTATCCAGGCCTCTGGGTGTCCCGACACAGGTG
TTCACATCTGTGCTGTGAGGTCAGATGCCTCAGTTCTTGAAAGCTAGGTTCTGCGACTGTTAC
CAAGGTGATTGTAAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGTGAA
TCGGACAGCCTCCAGCAGAGGTGTGGGAGCTGCAGCTGAGGGAAGAAGAGACAATCGGCCTGGA
CACTCAGGAGGGTCAAAGGAGACTTGGTCGACCACTCATCCTGCCACCCCAAGATGCATCCT
GCCTCATCAGGTCCAGATTTCTTTCCAAGGCGGACGTTTCTGTTGGAATTCTTAGTCTTGGCC
TCGGACACCTTCATTCTGTTAGCTGGGGAGTGGTGGTGAGGCAGTGAAGAAGAGGCGGATGGTCAC
ACTCAGATCCACAGAGCCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGGCCCCC
ACCCCAACCTGCACAGCCCTCATCCCTCTTGGCTTGAGCCGTCAGAGGCCCTGTGCTGAGTGT
CTGACCGAGACACTCACAGCTTGTGATCAGGGCACAGGCTTCTCGGAGCCAGGATGATCTGTG
CCACGCTTGCACCTCGGGCCCATCTGGGCTCATGCTCTCTCTCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTAGTTACCAAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTTLPPPLF
SKVVIVLIDALRDDDFVFGSKGVKFMPTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDSEVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGTTSFFVSDYTEVDNNV
TRHLDKVLKRGDWDILILHYLGLDGHIGHISGPNPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGGTTATCGTGACGCACCTTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGTTGGAGTACAAC
TCCCTATAGAAAACAACTGCCAGCACCTTAAGACCACCTCACACCTTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCAGCATTTCATGACCAGGATCACAAAGTACTGGTCCTGGACTCTGGGAATCTCAT
AGCAGTTCCAGATAAAAACCTACATACGCCAGAGATCTTCTTTGCATTAGCCTCATCCTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGGTCTCTAAAGGGGAGTTTGTCTCTAC
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAGAACTGATGAAGCT
GGCTGCCCCAAAAGGAATCAGCACGCCGGCCCTTCATCTTTTATAGGGCTCAGGTGGGCTCCTGGA
ACATGCTGGAGTCGGCGGCTCACCCCGGATGGTTTCATCTGCACCTCCTGCAATTGTAATGAGCCT
GTTGGGGTGACAGATAAATTTGAGAACAGGAAACACATTGAATTTTCATTTCAACCAGTTTGCAA
AGCTGAAATGAGCCCCAGTGAGGTCAGCGATTAGGAACTGCCCCATTGAACGCCCTTCCTCGCTA
ATTGAACTAATTGTATAAAAACACCAAACCTGCTCACT

FIGURE 142

MLLLLEYNFPIENNCQHLKTTHTFRVKNLNPCKFSIHDQDHKVLVLDSGNLIAVPDKNYIRPEI
FFALASSLSSASAEKGSPIILGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAQKESARRPFI
FYRAQVGSWNMLESAAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 33-36

N-myristoylation site.

amino acids 50-55, 87-92

Interleukin-1

amino acids 37-182

FIGURE 143

CTAGAGAGTATAGGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTTAA
TCCAGGATCCTGTCCTTCCTGTCCTGTAGGAGTGCCTGTTGCCAGTGTGGGGTGAGACAAGTTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATTAAGGGCTGGGTCTGTGCTCAATTAACCTGTGGG
CACGGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCAATGCTGGGCCTGCCGTGG
AAGGGAGGTCTGTCCTGGGCGCTGCTGCTGCTTCTCTTAGGCTCCCAGATCCTGCTGATCTATGC
CTGGCATTTCACGAGCAAAGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTTGTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCTACAGACTGGGG
CACATCTTGAATTCCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGGAAATTTGAAGACGACATTGACAACCTGCCATTTCCAAGAAAGCACAG
AGCTGAACAATACTTTACCTGCTTCTTCACCATCAGCACCAGGCCCTGGATGACTCAGTTCAGC
CTCCTGAACAAGACCTGCTTGGAGGGATTCCACTGAGTGAAACCCACTCACAGGCTTGTCCATGT
GCTGCTCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTTCAGTGGCTGAGCAGCT
TTGGACTTGTGTGTTATCCATTTTTGCATGTGTTTGAGATCTCAGATCAGTGTTTTAGAAAATCC
ACACATCTTGAGCCTAATCATGTAGTGTAGATCATTAACATCAGCATTTTAAGAAAAAAAAAAAA
AA

FIGURE 144

MLGLPWKGGLSWALLLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLNKTCLLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGTGGCTGGCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAATCAAGTGGAACCGGAAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGCCCAGGTGGCTGAGAACCGCCCGGGAGCCTTCATCAAGCAAGGCCGCAAGCTCGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTTGTACCGGCTGCATCAATGCCA
CCCAGGCGGCGAACCAGGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTTGGTTGGAGAGGGGCGCAGGACT
TCGGGTCACCATGCACCAGCCAGTGCTCCTCTGCCTTCTGGCTTTGATCTGGCTCATGGTGAAT
AAGCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAAATCCTGGCAAGTGACCCAGCT
CTTCTCCCCCAAACCCACGCGTGTCTGAAGGTGCCCAGGAGCGGCGATGCACTCGCACTGCAAA
TGCCGCTCCCACGTATGCGCCCTGGTATGTGCCTGCGTTCTGATAGATGGGGGACTGTGGCTTCT
CCGTCACTCCATTCTCAGCCCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTTCACAGTACTTCCCAACAACCTCTTAGAGGTAG
GTGTATTCCCGTTTTACAGATAAGGAAACTGAGGCCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCTGGCTTGTCTAACCCAGGTTTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTGATCACTTTATGTCTCACAGGGACCCACATCCAAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTTAGAAATAAATGAAAACACCTGA

FIGURE 146

MRKHLSSWWLATVCMLLFSHLSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNKLHQQ
VLWRLVQELCSLKHCEFWLERGAGLRVTMHQPVLLCLLALIWLMMVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCTGGTCCAGAGTCTCATT
CCTGATGATTTATAGACTCAAAGAAAACTATGTTTCAGAAGCTCTCTTCTCTTCTGGCCTCCTCT
CTGTCTTCTTTCCCTCTTTCTTCTTATTTTAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTTCATTTTGCTTGTGTCAGTGGGGTAGGTCAGTCTTAGTTTTATTTTTGAAATTT
CAACTTTCAGATTCAGGGGGTACATGTGAAGGTTTGTTTTATGAGTATATTGCATTGATGCTGAGG
TTTGGGGT

APP ID=10063599

FIGURE 148

MFRSLLFWPPLCLLSLFLILISSIYSESKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAACCTCAGACCCACAGGGCGGACAGCGCTCCCCTCTACCTGGAGACTTGAC
TCCCGCGCGCCCCAACCCCTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCCTGCAGCCGCCAGTCC
CGGCCCTCTCCCGCCCCACCCACCCCTCCTGGCTCTTCTGTTTTTACTCCTCCTTTTCATTCTATA
ACAAAAGCTACAGCTCCAGGAGCCAGCGCGGGCTGTGACCCAAGCCGAGCGTGAAGAATGGGGTT
CCTCGGGACCGGCACTTGGATTCTGGTGTAGTGCTCCCGATTCAAGCTTTCCCCAACCTGGAGGAA
GCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCAGAAAAGACCTTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAAAAAACATATCCTCCAGAAAACAAGCCAGGTGAGAGCAACTATTCTTT
TGTTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAGGAAAAAATTGAGAAAAGACAATCTA
TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAACG
ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCA
TCAACTAGACGGGACTCCTTTAACCGCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTATGAAG
AAAATGACAGAGCCGTGTTTGACAAGATTGTTTCTAACTACTTAATCTCGGCCTTATCACAGAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTTACAAAAATTAATCTCAAAGGAAGCCAACAA
TTATGAGGAGGATCCCAATAAGCCCACAAGCTGGACTGAGAATCAGGCTGGAAAAATACCAGAGAAAG
TGACTCCAATGGCAGCAATTCAGATGGTCTTGCTAAGGGAGAAAACGATGAAACAGTATCTAACACA
TTAACCTTGACAAATGGCTTGGAAGGAGAACTAAAACCTACAGTGAAGACAACCTTGAGGAACCTCA
ATATTTCCCAAATTTCTATGCGCTACTGAAAAGTATTGATTGAGAAAAGAAGCAAAAGAGAAAGAA
CACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAGATGATGGTGAAATATGGAACAATATCT
CCAGAAGAAGGTGTTTCTACCTTGAAAACCTGGATGAAATGATTGCTCTTCAGACCAAAAACAAGCT
AGAAAAAAATGCTACTGACAATATAAGCAAGCTTTTCCCAGCACCATCAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTTGGAAGC
CATCAGAAAAAATATTGAATGGTTGAAGAAACATGACAAAAAGGAAATAAAGAAGATTATGACCTTT
CAAAGATGAGAGACTTCATCAATAAACAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAA
GAAGCCGAGGCCATCAAGCGCATTTATAGCAGCCTGTAAAAATGGCAAAGATCCAGGAGTCTTTCAA
CTGTTTCAGAAAACATAATATAGCTTAAAAACACTTCTAATTCTGTGATTAAATTTTTTGACCCAAGG
GTTATTAGAAAGTGCTGAATTTACAGTAGTTAACCTTTTACAAGTGGTTAAACATAGCTTTCTTCCC
GTAAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLP IQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIPEKVTMAAIQDGLAKGENDETVSNTLTLTNGLE
RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNKLEKNATDNISKLFAPSEKSHEETDSTKEEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHKDKGNKEDYDLSKMRDFINKQADAYVEKGILDK
EEAEAIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCTGTGGAGCTCAA
GATGGTCTCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTTATCTGC
ATAATAACCAGCTTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCATTAAAGGTGAAGAGATCAGC
GTGGTCCCCAATCGGTGGTGGATGCCAGCTGTCCCCGTATCCTGGGTGTCCAGGGTGGAAAG
CCAGTGCCTGTCTGTGGGGTGGGGCAGGAGCCGACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTTGGTGCCAAGGAATCCAAGAGCTTCACCTTCTACCGCGGGACATGGGGCTCACC
TCCAGCTTCGAGTCGGCTGCCTACCGGGCTGGTTCTGTGCACGGTGCCTGAAGCCGATCAGCC
TGTGAGACTCAGCCAGCTTCCCGAGAATGGTGGCTGGAATGCCCCATCACAGACTTCTACTTCC
AGCAGTGTGACTAGGGCAACGTGCCCCCAGAACTCCCTGGGCAGAGCCAGCTCGGGTGGGGGT
GAGTGGAGGAGACCCATGGCGGACAATCACTCTCTGTCTCAGGACCCCCACGTCTGACTTAG
TGGGCACCTGACCACTTTGTCTTCTGGTTCCAGTTTGGATAAATTCTGAGATTTGGAGCTCAGT
CCACGGTCTCCCCACTGGATGGTGTCTACTGTGTGGAACCTTGTAAAAACCATGTGGGGTAA
CTGGGAATAACATGAAAAGATTCTGTGGGGGTGGGGTGGGGGAGTGGTGGGAATCATTCTGTCT
TAATGGTAAGTACAAGTGTACCTGAGCCCCGAGGCCAACCCATCCCCAGTTGAGCCTTATA
GGGTGAGTAGCTCTCCACATGAAGTCTGTCACTCACCCTGTGCAGGAGAGGGAGGTGGTCATA
GAGTCAGGGATCTATGGCCCTTGGCCAGCCCCACCCCTTCCCTTTAATCCTGCCACTGTCTATA
TGCTACCTTTCTATCTCTTCCCTCATCATCTTGTGTGGGCATGAGGAGGTGGTGTGATGTCAGAA
GAAATGGCTCGAGCTCAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTCTTTTAAAAACCCAA
GATACAATCAAAATCCCAGATGCTGGTCTCTATTCCCATGAAAAAGTGTCTATGACATATTGAGA
AGACCTACTTACAAAGTGGCATATATTGCAATTTATTTTAAATTAAGATACCTATTTATATATT
TCTTTATAGAAAAAGTCTGGAAGAGTTTACTTCAATTGTAGCAATGTGAGGGTGGTGGCAGTAT
AGGTGATTTTTCTTTAATTCTGTTAATTTATCTGTATTTCTAATTTTTCTACAATGAAGATGA
ATTCTTGTATAAAAAATAAGAAAAGAAATTAATCTTGAGGTAAAGCAGAGCAGACATCATCTCTGA
TTGTCTCAGCCTCCACTTCCCCAGAGTAAATTCAAATTTGAATCGAGCTCTGCTGCTCTGGTTGG
TTGTAGTAGTGATCAGGAAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTTGT
GTGGCTGGAATCTCTGGGTAAGGAACCTAAAGAACAATAATCATCTGGTAATTCTTTCTAGAAAG
GATCACAGCCCCCTGGGATTCCAAGGCATTGGATCCAGTCTCTAAGAAGGCTGTGCTGACTGGTTGA
ATTGTGTCCCCCTCAAATTCACATCCTTCTTGGAAATCTCAGTCTGTGAGTTTATTTGGAGATAAG
GTCTCTGCAGATGTAGTTAGTTAAGACAAGGTCATGCTGGATGAAGGTAGACCTAAATTCATAT
GACTGGTTTCTTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCGGGGAAGACTATGTA
AAGATGAAGGCAGAGATCGGAGTTTTGCAGCCACAAGCTAAGAAACACCAAGGATTGTGGCAACC
ATCAGAAGCTTGGAAAGAGGCAAGAAGAATTCTTCCCTAGAGGCTTTAGAGGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCCTGAACGAAGAAAGAATAAATTTCCGGCTGTTTTAA
GCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAAACTAATACAGCTGCTAAATGATCCCTGT
CTCCTCGTGTTTACATTCTGTGTGTGTCCTTCCACAATGTACCAAAGTTGTCTTTGTGACCAA
TAGAATATGGCAGAAGTGTGGCATGCCACTTCCAAGATTAGGTTATAAAAGACACTGCAGCTTTC
TACTTGAGCCCTCTCTCTGTGCCACCCACCGCCCCAATCTATCTTGGCTCACTCGCTCTGGGGG
AAGCTAGCTGCCATGCTATGAGCAGGCCTATAAAGAGACTTACGTGGTAAAAAATGAGTCTCCT
GCCCCAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATGTTTGTGTTTT
AAGTTGCTCAGTTTTGGTCTAACTTGTATGCAGCAATAGATAAATAATATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEIISVVPNRWLDASLSPVILGVQGG
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGGWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTCTTTCCTTATGGGGACCTGGCCACCAGCTGCCTCCTTCTCTTGG
CCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
CAACACAGACGTTTCGTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACTTCACCCTTGAAGAAGTGTGTTCCCTCAATCTGATAGGTTC
CAGCCTTATATGCAGGAGGTGGTGCCCTTCCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTGCTGTTTATGTCTCTGAGAAAT
GCCTGCATTGACCAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTTCCTGCTAGAAATAA
CAATTAGATGCCCCAAAGCGATTTTTTTTAAACAAAAGGAAGATGGGAAGCCAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAACCAATGCCACTTTTGTTTATA
AGACCAGAAGGTAGACTTTCCTAAGCATAGATATTTATTGATAACATTCATTGTAAGTGGTGTTT
TATACACAGAAAACAATTTATTTTTTAAATAATTGTCTTTTCCATAAAAAAGATTACTTTCCAT
TCCTTTAGGGGAAAAAACCCCTAAATAGCTTCATGTTTCCATAATCAGTACTTTATATTTATAAA
TGTATTTATTATTATTATAAGACTGCATTTTATTTATATCATTTTATTAATATGGATTTATTTAT
AGAAACATCATTCGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTTATGACAATAATTAT
AGAGCTATAACATGTTTATTTGACCTCAATAAACACTTGGATATCCC

FIGURE 154

MAALQKSVSSFILMGTLATSCLLLLLALLVQGGAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNTDVRLIGEKLFHGVSMSERCYLMKQVLNFTLEEVLFPQSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
CAGTCAGTGCCCGACTTGTGACTGAGTGTGCAGTGCCCGAGCATGTACCAGGTCAGTGCAGAGGGC
TGCCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTTTGGGGCTGGGGGCCAAGTGGAGTGAGAACTGGGATCCCAGGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTTTCCTACAGGTGGTTGCAT
TCTTGGCAATGGTCATGGGAACCCACACCTACAGCCACTGGCCCAGCTGCTGCCCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCTCCCCTAGAGCCTGTAG
GCCCCAACCGCCACCCAGAGTCTGTAGGGCCAGTGAAGATGGACCCCTCAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCAGGACCTGTACCACGCCCCGT
TGCTGTGCCCCGACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCCGGGGCAACTCGGA
GCTGCTCTACCACAACCAGACTGTCTTCTACAGGCGGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTTAGCTTGTGTGTGTGTGCGGCCCCGT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTTGGGAAACCTGGAGCCAGGTGTACA
ACCACTTGCCATGAAGGGCCAGGATGCCCAGATGCTTGCCCCCTGTGAAGTGTGTCTGGAGCAG
CAGGATCCCGGGACAGGATGGGGGGCTTTGGGGAAAACCTGCACTTCTGCACATTTTGAAAAGAG
CAGCTGCTGCTTAGGGCCGCCGAAGCTGGTGTCTGTCTCTCTCAGGAAAGGTTTTCAA
GTTCTGCCCATTCTGGAGGCCACCACTCCTGTCTCTCTCTCTTTTCCCATCCCCTGCTACCCCTG
GCCCAGCACAGGCACTTTCTAGATATTTCCCCCTTGCTGGAGAAGAAAGAGCCCCCTGGTTTTATT
TGTTTGTTTACTCATCACTCAGTGAGCATCTACTTTGGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTTATCCAAATAAATAT
CTTTATTTAAAAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGHTYSHWPSCCPKSGQDTSEELLRWSTVPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTGKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCGATGTCGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAGCC
GACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACAGTGTGCAACAGGGGACTATTCA
ATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTTGAAGGCCACCAA
GATTTGTGTGACGGGCAAAGCAACTTCCAGTCCTACAGCTGTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATGGACATTTTCCTACATCGGCTTCCCTGTA
GAGCTGAACACAGTCTATTTTATTGGGGCCATAATATTCCTAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTTACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAA
AGTGTGTCAAGGCCGAAGCCTGTGGGATCCGAACATCACTGCTTGTGAAGAAGATGAGGAGACA
GTAGAAGTGAACCTTCAACCACTCCCTGGGAAACAGATACATGGCTCTTATCCAACACAGCAC
TATCATCGGGTTTTCTCAGGTGTTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGACGCTGACTCCATATTTTCCTACTTGTGGC
AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCCAACAGGCGTCCCTTTCCCTCT
GGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCCTGCTGTCTCTGCTGGTGGCCA
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCCTTT
TCTACCACCACACTACTGCCCCCATTAAAGTTCCTTGTGGTTTACCCATCTGAAATATGTTTCCA
TCACACAATTTGTTACTTCACTGAATTTCTTCAAAACCATTGCAGAAGTGAGGTATCCTTGAAA
AGTGGCAGAAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTTGCCACTCAAAGAAGGCA
GCAGACAAAGTCGTCTTCTTCTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCAGTGAGAACTCTCAAGACCTCTTCCCCCTTGCCCTTTAACCTTTTCTGCA
GTGATCTAAGAAGCCAGATTCTCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTT
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCTCCTTGTAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLPGLRDLRVEPVTTSVATGDYSILMNVSWV
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP
NANMNEDGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKNEETVEVNFTTTPLGNRYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPVTGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPTGVPFPLDNNK
SKPGGWLPLLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTTLLPPIKVLVVYPSEICFHHTICYFTEFL
QNHCRSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLLSNDVNSVCDGTCGKSEGSPSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVC PKYHLMKDATAFCAELLHV KQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGCAACATGACAGTGAAGACCCTGCATGGCCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGGCTTGCCTTTCTGAGTGAGGCGGCAGCTCGGAAAATCCCCAAAGTAGGACATA
CTTTTTTCCAAAAGCCTGAGAGTTGCCCCGCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCGTTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG
GAATTACACTGTCACTTGGGACCCCAACCGGTACCCCTCGGAAGTTGTACAGGCCCAAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCCATCCAGCAA
GAGACCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTTCTTCCAGTTGGAGAAGGTGCT
GGTGACTGTTGGCTGCACCTGCGTCAACCCTGTCATCCACCATGTGCAGTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 160

MTVKTLHGPMVKYLLLSILGLAFLSEAAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIINEN
QRVSMSRNIESRSTSPWNYTWTDPNRYPSEVVQAQCRNLGCINAQGKEDISMNSVPIQQETLVV
RRKHQGCSVSFQLEKVLVTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

A C A C T G G C C A A A C A A A A C G A A A G C A C T C C G T G C T G G A A G T A G G A G G A G A G T C A G G A C T C C C A G G
 A C A G A G A G T G C A C A A A C T A C C C A G C A C A G C C C C C T C C G C C C C C T C T G G A G G C T G A A G A G G G A T T C
 C A G C C C C T G C C A C C C A C A G A C A C G G G C T G A C T G G G G T G T C T G C C C C C C T T G G G G G G G G C A G C A C
 A G G G C C T C A G G C C T G G G T G C C A C C T G G C A C T A G A A G A T G C C T G T G C C C T G G T T C T T G C T G T C C T
 T G G C A C T G G G C C G A A G C C C A G T G G T C C T T T C T C T G G A G A G G C T T G T G G G C C T C A G G A C G C T A C C
 C A C T G C T C T C C G G G C C T C T C C T G C C G C C T C T G G G A C A G T G A C A T A C T C T G C C T G C C T G G G G A C A T
 C G T G C C T G C T C C G G G C C C C G T G C T G G C G C C T A C G C A C C T G C A G A C A G A G C T G G T G C T G A G G T G C C
 A G A A G G A G A C C G A C T G T G A C C T C T G T C T G C G T G T G G C T G T C C A C T T G G C C G T G C A T G G G C A C T G G
 G A A G A G C C T G A A G A T G A G G A A A A G T T T G G A G G A G C A G C T G A C T C A G G G G T G G A G G A G C C T A G G A A
 T G C C T C T C T C C A G G C C C A A G T C G T G C T C T C C T T C C A G G C C T A C C C T A C T G C C C G C T G C G T C C T G C
 T G G A G G T G C A A G T G C C T G C C C T T G T G C A G T T T G G T C A G T C T G T G G G C T C T G T G G T A T A T G A C
 T G C T T C G A G G C T G C C C T A G G G A G T G A G G T A C G A A T C T G G T C C T A T A C T C A G C C C A G G T A C G A G A A
 G G A A C T C A A C C A C A C A C A G C A G C T G C C T G C C C T G C C C T G G C T C A A C G T G T C A G C A G A T G G T G A C A
 A C G T G C A T C T G G T T C T G A A T G T C T C T G A G G A G C A G C A C T T C G G C C T C T C C C T G T A C T G G A A T C A G
 G T C C A G G G C C C C C A A A A C C C G G T G G C A C A A A A C C T G A C T G G A C C G C A G A T C A T T A C C T T G A A
 C C A C A C A G A C C T G G T T C C C T G C C T C T G T A T T C A G G T G T G G C C T C T G G A A C C T G A C T C C G T T A G G A
 C G A A C A T C T G C C C C T T C A G G A G A C C C C G C G C A C A C C A G A A C C T C T G G C A A G C C G C C C G A C T G
 C G A C T G C T G A C C C T G C A G A G C T G G C T G C T G G A C G C A C C G T G C T C G C T G C C C G C A G A A G C G G C A C T
 G T G C T G G C G G G C T C C G G G T G G G A C C C C T G C C A G C C A C T G G T C C C A C C G C T T T C C T G G G A G A A C G
 T C A C T G T G G A C A A G G T T C T C G A G T T C C C A T T G C T G A A A G G C C A C C C T A A C C T C T G T G T T C A G G T G
 A A C A G C T C G G A G A A G C T G C A G C T G C A G G A G T G C T T G T G G G C T G A C T C C C T G G G G C C T C T C A A A G A
 C G A T G T G C T A C T G T T G G A G A C A C G A G G C C C C A G G A C A A C A G A T C C C T C T G T G C C T T G G A A C C C A
 G T G G C T G T A C T T C A C T A C C C A G C A A A G C C T C C A C G A G G G C A G C T C G C C T T G G A G A T A C T T A C T A
 C A A G A C C T G C A G T C A G G C C A G T G T C T G C A G C T A T G G G A C G A T G A C T T G G G A G C G C T A T G G G C C T G
 C C C C A T G G A C A A A T A C A T C C A C A A G C G C T G G G C C C T C G T G T G G C T G G C C T G C C T A C T C T T T G C C G
 C T G C G C T T T C C C T C A T C C T C C T T C T C A A A A A G G A T C A C G C G A A A G G G T G G C T G A G G C T C T T G A A A
 C A G G A C G T C C G C T C G G G G G C G C C G C C A G G G C C G C G C G C T C T G C T C C T C T A C T C A G C C G A T G A
 C T C G G G T T T C G A G C G C C T G G T G G G C G C C C T G G C G T C G G C C C T G T G C C A G C T G C C G T G C G C G T G G
 C C G T A G A C C T G T G G A G C C G T C G T G A A C T G A G C G C G C A G G G C C C G T G G C T T G G T T T C A C G C G C A G
 C G G C G C C A G A C C C T G C A G A G G G C G G C G T G G T G G T C T T G C T C T C T C C C G G T G C G G T G G C G C T
 G T G C A G C G A G T G G C T A C A G G A T G G G G T G T C C G G G C C C G G G G C G C A C G C C C G C A C G C C T T C C
 G C G C C T C G C T C A G C T G C G T G C T G C C C G A C T T C T T G C A G G G C C G G G C G C C C G G C A G C T A C G T G G G G
 G C C T G C T T C G A C A G G C T G C T C A C C C G G A C G C C G T A C C C G C C C T T T T C C G C A C C G T G C C C G T C T T
 C A C A C T G C C C T C C C A A C T G C C A G A C T T C C T G G G G G C C C T G C A G C A G C C T C G C G C C C G C G T T C C G
 G G C G G C T C C A A G A G A G A G C G G A G C A A G T G T C C C G G G C C C T T A G C C A G C C C T G G A T A G C T A C T T C
 C A T C C C C G G G G A C T C C C G C C C G G A C G C G G G T G G G A C C A G G G C G G G A C C T G G G G C G G G G A
 C G G G A C T T A A A T A A A G G C A G A C G C T G T T T T C T A A A A A A

FIGURE 162

MPVPWFLLSLALGRSPVLSLERLVGPQDATHCSPGLSCRLWSDILCLPGDIVPAPGPVLAPTHLQTELV
LRCQKETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNASLQAQVLSFQAYPTARCVLLEV
QVPAALVQFGQSVGSVVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNVSDGDNVHLVINVS
EEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQN
LWQAARLRLLTLQSWLLDAPCSLPAAEALCWRAPGGDPCQPLVPPLSWENVTVDKVLEFPLLKGHPNLCVQ
VNSSEKLQIQECLWADSLGPKDDVLLLETRGPQDNRSLEPSGCTSLPSKASTRAARLGEYLLQDLQS
GQCLQLWDDDLGALWACPMCKYIHKRWALVWLACLLFAAALSLILLKKDHAKGWLRLKQDVRSGAAARG
RAALLLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAVWFHAQRRQTLQEGGVVLLFSP
GAVALCSEWLQDGVSGPGAGHPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLHHPDAVPALFRTVPVFT
LPSQLPDFLGALQQPRAPRSGRLQERAEQVSRAEQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTTGACTGTGGGATCCCTGGCT
 GCTCACGCCCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCAGTCCAGCAACTTTGA
 AAACATCCTGACGTGGGACAGCGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA
 AGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTCAGCGGATCACCCGGAAGTCTGC
 AACCTGACGGTGGAGACGGGCACCTCACGGAGCTCTACTATGCCAGGGTCACCGCT
 GTCAGTGGGGAGGCCGGTCAGCCACCAAGATGACTGACAGGTTAGCTCTCTGCAGCACACTAC
 CCTCAAGCCACCTGATGTGACCTGTATCTCCAAAGTGAGATCGATTGATGATTGTTTCATCCTA
 CCCCCACGCCAATCCGTGCAGGCGATGGCCACCGGCTAACCTGGAAGACATCTTCCATGACCTG
 TTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAATGACCTTGGAGGGAAGCAGAGAGA
 ATATGAGTTCTTCGGCCTGACCCCTGACACAGATTCTTGGCACCATCATGATTGCGTTCCCA
 CCTGGGCCAAGGAGAGTGGCCCTACATGTGCCGAGTGAAGACACTGCCAGACCGGACATGGACC
 TACTCCTTCTCCGGAGCCTTCTGTTCTCCATGGGCTTCTCGTCGAGTACTCTGCTACCTGAG
 CTACAGATATGTACCAAGCCGCTGCACCTCCCAACTCCCTGAACGTCCAGCGAGTCTGACTT
 TCCAGCCGCTGCGCTTCATCCAGGAGCACGTCTGATCCCTGTCTTTGACCTCAGCGGCCCCAGC
 AGTCTGGCCCCAGCCTGTCCAGTACTCCCAGATCAGGGTGTCTGGACCCAGGGAGCCCGCAGGAGC
 TCCACAGCGGCATAGCCTGTCCGAGATCACCTACTTAGGGCAGCCAGACATCTCCATCCTCCAGC
 CCTCCAACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCTATGCCCCAAACGCTGCCCTGAG
 GTCGGGCCCCCATCCTATGCACCTCAGGTGACCCCCGAAGCTCAATTCCATTCTACGCCCCACA
 GGCCATCTCTAAGGTCCAGCCTTCTCCTATGCCCCCTCAAGCCACTCCGGACAGCTGGCCTCCCT
 CCTATGGGGTATGCATGGAAGTTCTGGCAAAGACTCCCCACTGGGACACTTTCTAGTCTCTAA
 CACCTTAGGCCATAAGGTGACCTTTCAGAAAGAGCCACCAGCTGGAAGCTGCATGTTAGGTGGCCT
 TTCTCTGCAGGAGGTGACCTCCTTGGCTATGGAGGAATCCCAAGAAGCAAAATCATTCGACCAGC
 CCCTGGGGATTGACAGACAGAATCTGACCCAAATGTGCTACACAGTGGGGAGGAAGGGACA
 CCACAGTACCTAAAGGGGCCAGCTCCCCCTCCTCTCCTCAGTCCAGATCGAGGGCCACCCCATGTC
 CCTCCCTTTGCAACCTCCTTCCGGTCCATGTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGGCC
 TGCTGGAGTCCCTTGTGTGTCCCAAGGATGAAGCCAAGAGCCAGCCCTGAGACCTCAGACCTG
 GAGCAGCCACAGAATGGATCTCTTTTTCAGAGGCCCTGGCCCTGACTGTGCAGTGGGAGTCTG
 AGGGGAATGGGAAAGGCTTGGTGCTTCCCTCCCTGTCCCTACCCAGTGTACATCCTTGGCTGTCA
 ATCCCATGCCCTGCCATGCCACACACTCTGCGATCTGGCCTCAGACGGGTGCCCTTGGAGAAAGC
 AGAGGGAGTGGCATGCCAGGGCCCTGCCATGGGTGCGCTCCTCACCGGAACAAAGCAGCATGATA
 AGGACTGCAGCGGGGAGCTTGGGGAGCAGCTTGTGTAGACAAGCGCGTGCTCGCTGAGCCCTG
 CAAGGCAGAAATGACAGTGAAGGAGGAATGCAGGGAAACTCCCGAGGTCCAGAGCCCCACCTC
 CTAACACCATGGATTCAAAGTGCTCAGGGAATTTGCCTCTCCTTGGCCCATTCCTGGCCAGTTTC
 ACAATCTAGCTCGACAGAGCATGAGGCCCTGCCTCTTCTGTCTATTGTTCAAAGGTGGGAAGAGA
 GCCTGGAAAAGAACAGGCCCTGGAAAAGAACAGAGGAGGCTGGGCAGAACAGAACAACTGC
 ACTTCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCTGCAGTCA
 TTCCCAGCCAGGGCAACTGCCTGACGTTGCACGATTTGAGCTTCATTCTCTGATAGAACAAAGC
 GAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCCTTTTCTGCAGGCAGGAGTTTCAGACCT
 ATCCTGAGAATGGGGTTTGAAGGAAGGTGAGGGCTGTGGCCCTGGACGGGTACAATAACACAC
 TGTACTGATGTCAAACTTTGCAAGCTCTGCCTTGGGTTGAGCCCATCTGGGCTCAAATTCAGC
 CTCACCACTCAAAAGCTGTGTGACTTCAAACAAATGAAATCAGTGCCAGAACCTCGGTTTCCTC
 ATCTGTAATGTGGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGTCATG
 TCTTTAAAGTGCTTAATAGTGCCTGGTACATGGGCAGTGCCCAATAAACGGTAGCTATTTAAAAA
 AAAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGERDW
VAKKGCQRITRKSCNLTVETGNLTelyyArvtAVSAGGRSATKMTDRFSSLQHTTLKPPDVTCIS
KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCrvKTLpDRTWTYSFSGAFLESMGFLVAVLCYLSYRYVTKPPAP
PNSLNVQRVLTFQPLRFIQEHVLIPVFDLSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFFFYAPQAISKVQPSSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLSspKHLRPKGQLQKEPPAGSCMLGGLSLQEVTSLAM
EESQEAKSLHQPLGICTDRtSDPNVLHSGEEGTPQYLKGQLPLLSSVQIEGHPMSLPLQPPSGPC
SPSDQGPSPWGLLLESLVCPKDEAKSPAPETSdLEQPTELDSLFRGLALTvQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAAAAAAGTACCCGGGCGCGGTGGCCACAACATGG
CTGCGGCGCGGGGGCTGCTCTTCTGGCTGTTTCGTGCTGGGGGCGCTCTGGTGGGTCCCGGGCCAG
TCGGATCTCAGCCACGGACGGCGTTTCTCGGACCTCAAAGTGTGCGGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGGAAAGCTCTTGAAGACTTCACGGGCCCTGATTGTCGTTTTGTGAATTTTA
AAAAAGGTGACGATGTATATGTCTACTACAACTGGCAGGGGGATCCCTTGAACCTTGGGCTGGA
AGTGTGAACACAGTTTTGGATATTTCCAAAAGATTTGATCAAGGTACTTCATAAATACACGGA
AGAAGAGCTACATATTCAGCAGATGAGACAGACTTTGTCTGCTTTGAAGGAGGAAGAGATGATT
TTAATAGTTATAATGTAGAAGAGCTTTTAGGATCTTTGGAACTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGA
ACTTGACCCTGTGCTGAGCCCGAGGCATTCAGAGCTGATTCAGAGGATGGAGAAGGTGCTTTCT
CAGAGAGCACCGAGGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCTCACACCAGCGGT
CCTGCGGCTAACGCTCAGGGAGTGCAGTCTTCGTTGGACACTTTTGAAGAAATTCGCACGATAA
ATTGAAAGTGCCGGGAAGCGAAAGCAGAAGTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCCTTATT
CATTACAGCAAAGGATTTTCGTTGGCATCAAATCTAAGTTTGTTTTACAAAGATTGTTTTTAGTA
CTAAGCTGCCTTGGCAGTTTGCATTTTTGAGCCAAACAAAATATATTATTTCCCTTCTAAGTA
AAAAAAAAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFLGALWWVPGQSDLHGRRFSDLKVCGDEECSMLMYRGKALEDFTGPDCRFVN
FKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEELHIPAETDFVCFEGGRD
DFNSYNVEELLGSLELEDSPVEESKKAEEVSQHREKSPEESRGRELDPVPEPEAFRADESDEGEA
FSESTGLQGQPSAQESHPTSGPAANAQGVQSSLDTFEEILHDKLKVPGESRTGNSSPASVER
EKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNLSLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTGAGAGGCCGGGGAAGAGAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCTCTAGGACATACACGGGACCCCTAACTTC
AGTCCCCCAAACGCGCACCCCTCGAAGTCTTGAAGTCCAGCCCCGCACATCCACGCGGGCACAGG
CGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTTCGGGCAGCTGGGCTCGGGC
GGCGGGAGTAGGGCCCGGCAGGGAGGCAGGGAGGCTGCATATTCAGAGTCGCGGGCTGCGCCCTG
GGCAGAGGCCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCGATGAGCCGCGTGG
TCTCGCTGCTGCTGGGCGCCGCGCTGCTCTGCGGCCACGGAGCCTTCTGCCGCCGCGTGGTCAGC
GGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGGCCTACTTCCATGAAGT
GTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCTGGCTTGTGAGAGTGAGGGAGGAGTCCCTCCTCA
GCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAACCTGACAAAACCCGGG
ACAGGGATTTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGAATGGAGATGGGCAAACATCTGG
TGCCTGCCCAGATCTCTACAGTGGTCTGATGGAAGCAATTCAGTACCGAACTGGTACACAG
ATGAACCTTCCTGCGGAAGTGAAAAGTGTTGTGATGTATCACCACCAACTGCCAATCCTGGC
CTTGGGGGTCCCTACCTTTACCAGTGGAATGATGACAGGTGTAACATGAAGCACAATTATATTTG
CAAGTATGAACCAGAGATTAATCCAACAGCCCTGTAGAAAAGCCTTATCTTACAATCAACCAG
GAGACACCCATCAGAATGTGGTTGTTACTGAAGCAGGTATAATTCCCAATCTAATTTATGTTGTT
ATACCAACAATACCCCTGCTCTTACTGATACTGGTTGCTTTTGAACCTGTTGTTTCCAGATGCT
GCATAAAAGTAAAGGAAGAACAAAACTAGTCCAAACCAGTCTACACTGTGGATTTCAAAGAGTA
CCAGAAAAGAAAGTGGCATGGAAGTATAAATACTCATTGACTTGGTTCCAGAATTTTGTAATCT
GGATCTGTATAAGGAATGGCATCAGAACAATAGCTTGAATGGCTTGAAATCACAAGGATCTGC
AAGATGAACTGTAAGCTCCCCCTTGAGGCAAATATTAAAGTAATTTTTATATGCTATTATTTCA
TTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGCTAAAGGATGCACCCAA
ACTTCAAACCTCAAGCAAATGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTCGGGAGTA
TGTGTGTTAGAAGCAATTCCTTTTATTTCTTTACCTTTTCATAAGTTGTTATCTAGTCAATGTAA
TGTATATTGTATTGAAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGTGTTGATAAAA
ATGAACTGTTCTAATATTTATTTTATGGCATCTCATTTTTCAATACATGCTCTTTTGATTAAAG
AACTTATTACTGTTGTCAACTGAATTCACACACACACAAATATAGTACCATAGAAAAAGTTTGT
TTTCTCGAAATAATTCATCTTTTCAGCTTCTCTGCTTTTGGTCAATGTCTAGGAAATCTCTCAGA
AATAAGAAGCTATTTTCATTAAGTGTGATATAAACCTCCTCAAACATTTTACTTAGAGGCAAGGAT
TGTCTAATTTCAATTGTGCAAGACATGTGCCTTATAATTATTTTGTAGCTTAAAATTAACAGATT
TTGTAATAATGTAACCTTTGTTAATAGGTGCATAAACACTAATGCAGTCAATTTGAACAAAAGAAG
TGACATACACAATATAAATCATATGTCTTCACACGTTGCCTATATAATGAGAAGCAGCTCTCTGA
GGGTCTGAAATCAATGTGGTCCCTCTCTTGCCCACTAAACAAAGATGGTTGTTTCGGGGTTTGGG
ATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGCCTCTG
ACTATATTAGTATACAAAGAGGTCATGTGGTTGAGACCAGGTGAATAGTCACTATCAGTGTGGAG
ACAAGCACAGCACACAGACATTTTAGGAAGGAAAGGAAGTACGAAATCGTGTGAAATGGGTTGG
AACCCATCAGTGATCGCATATTCATTGATGAGGGTTTGTCTGAGATAGAAAATGGTGGCTCCTTT
CTGTCTTATCTCCTAGTTTCTTCAATGCTTACGCCTTGTCTTCTCAAGAGAAAGTTGTAAGTCT
CTGGTCTTCATATGTCCCTGTGCTCCTTTTAAACCAATAAAGAGTTCTTGTCTTGGGGGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSLLLGAALLCGHGAFRRVVSQGKVCFADFKHPCYK MAYFHELSSRVSFQEARLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDGFWIGLWRNGDGQTSGACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKS KGR TKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217